Decision Support Systems 209 (2026) 114740

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Beyond blind feature injection: An information foraging theory-guided deep learning framework for product recommendation

Weiyue Li [a,1], Ming Gao [a,1,*], Jingmin An [a,1], Bowei Chen [b,1], Jiafu Tang [a,1], Xuhui Wang [c,1], Yeming Gong [d,1]

[a] School of Management Science and Engineering, Key Laboratory of Big Data Management Optimization and Decision of Liaoning Province, Dongbei University of Finance and Economics, China
[b] Adam Smith Business School, University of Glasgow, United Kingdom
[c] School of Business Administration, Dongbei University of Finance and Economics, China
[d] Artificial Intelligence in Management Institute, EMLYON Business School, France

ARTICLE INFO

Keywords:
Recommender systems
Information foraging theory
Cognitive process-aware decision support
Deep learning
Design science

ABSTRACT

Recommender systems significantly improve the quality and confidence of consumer decisions by alleviating information overload through streamlined retrieval, playing a vital role in e-commerce platforms. However, most recommender systems underutilize contextual information and fail to explicitly model users' cognitive processes, leading to biased preference learning. Drawing on information foraging theory, this study adopts the design science research methodology to develop Tiresias, a framework that aligns with users' cognitive processes by mimicking how proximal cues and foraging costs are assessed for each product. Proximal cues such as product category and affinitive relationships, and foraging costs such as behavior type, dwell time, and foraging progression, are identified and encoded as learnable features. To model the product-foraging journey, Typhon is designed as a Kolmogorov–Arnold network-enhanced differential transformer with a non-invasive feature injection mechanism. A multi-task learning objective that combines next-product prediction with auxiliary cue and cost prediction refines contextualized preference representations. Experiments on datasets show that Tiresias achieves state-of-the-art performance, outperforming baselines by up to 5.44%. Finally, post-hoc analysis reveals learned representation patterns consistent with dynamic preference evolution throughout the user's product foraging journey. Methodologically, the study demonstrates how information foraging theory can guide feature selection, feature learning, and objective design. Practically, the study offers actionable insights for developing cognitive process-aware recommender systems that better support consumer decision-making.

1. Introduction

Recommender systems are core infrastructure in e-commerce platforms, driving user engagement, conversions, and substantial business value [1]. However, their effectiveness is constrained by data sparsity, arising from the mismatch between limited user attention and the scale of available products, which degrades recommendation performance [2–4]. To mitigate this issue, recent research incorporates side information to enrich user and item representations across five major paradigms: sequential [2], temporal [5], multi-behavior [3], attribute-aware [4], and social recommendation [6]. These approaches encode complementary signals, including interaction order and timing, heterogeneous user actions, item attributes, and user–user relationships. However, feature selection is typically driven by data availability rather than a principled understanding of how signals inform user decision-making, leading to the omission of meaningful cues and an incomplete representation of users' cognitive processes. For example, many social recommender systems [7–9] overlook dwell time and thus fail to capture users' depth of engagement with specific products.

Motivated by these limitations, this study examines two research questions: (i) which features should be selected from recommendation contexts, and (ii) how contextualized user preferences can be captured from these features. For the first, a promising approach is to ground feature selection in human decision-making processes. Cognitive theories provide guidance for decision support system design by clarifying how factors shape users' information processing [10], and their application in recommender systems has shown early success [11]. For instance,

* Correspondence to: No. 217 JianShan St., Shahekou District, Dalian, China.
E-mail address: gm@dufe.edu.cn (M. Gao).

1 These authors contributed equally to this work.

https://doi.org/10.1016/j.dss.2026.114740
Received 26 November 2025; Received in revised form 14 July 2026; Accepted 18 July 2026
Available online 24 July 2026

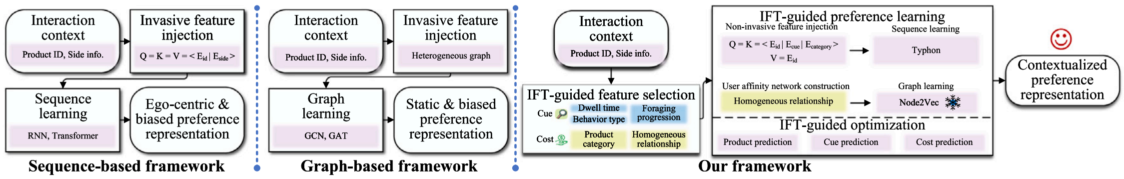


Fig. 1. Comparison of the proposed recommendation framework with existing approaches.

DeepRM-TC [11] aligns its architecture with four cognitive stages (i.e., interpretation, retrieval, judgment, and response selection) but focuses on modular design rather than identifying critical contextual features. For the second question, sequence-based and graph-based feature learning frameworks [12,13] are constrained by ineffective feature utilization and partial feature perception. As shown in Fig. 1, sequence-based methods fuse features invasively before sequential modeling, while graph-based methods encode them as nodes and edges in a heterogeneous graph. Both rely on auxiliary feature injection, where rich side information can overwhelm core product features [14], leading to unstable learning. They also capture features incompletely: sequence-based frameworks overlook affinitive user influence, whereas graph-based frameworks underutilize fine-grained temporal signals, resulting in suboptimal performance.

Following a design science research methodology [15], this study develops an artifact, Tiresias, to address the aforementioned two research questions. Tiresias is a cognitively grounded recommendation framework that leverages information foraging theory (IFT) to guide feature selection, feature learning, and parameter optimization. IFT conceptualizes navigation in complex environments as movement across information patches based on perceived value signals and foraging costs. In e-commerce, the product-seeking journey is modeled as an information foraging process in which users evaluate product pages by weighing proximal cues (e.g., product category and affinitive relationships) against foraging costs (e.g., behavior type, dwell time, and foraging progression). These cues and costs are parameterized as learnable feature representations to inform preference learning. Given the sequential and noisy nature of user interactions, the IFT-guided user preference extractor is designed as a denoising sequential model that captures the evolution of user preferences. Its core module, Typhon, employs a Kolmogorov–Arnold network-enhanced differential transformer with a non-invasive feature injection mechanism, enabling theory-guided features to be integrated without overwhelming primary product signals. To incorporate graph-structured affinitive relationships, a lightweight graph learning module is introduced to refine preference representations. A multi-task learning objective combines next-product prediction with auxiliary cue and cost prediction, encouraging the model to capture underlying decision mechanisms. Finally, the evaluation follows the design science validity framework [16] to articulate the study's knowledge claims.

This study makes three main contributions. First, it positions IFT as a key design principle for recommender systems, thereby contributing to the design science literature on theory-driven decision support. In the proposed framework, IFT implies that proximal cues and foraging costs guide the assessment of information patches, which in turn motivates a theory-based feature selection strategy, a multi-task learning objective, and the sequential modeling of the product-foraging journey implemented in Typhon. Second, the study broadens the application scope of IFT by using it to guide recommender system design. Although IFT has been employed to explain decision-making in information retrieval and to inform the design of software engineering tools [17,18] and user interfaces [19,20], its potential in recommender systems has yet to be fully realized; this work provides a concrete instantiation in this domain and offers a novel theoretical lens for understanding and modeling product-foraging sequences. Third, by adopting IFT as a foundational principle for structuring contextual signals, the framework shows that several seemingly distinct recommendation tasks, such as temporal and attribute-aware recommendation, can be interpreted as special cases within a unified formulation, thereby fostering a more integrated understanding of disparate recommendation settings. Extensive experiments on two real-world datasets demonstrate that Tiresias consistently outperforms 30 representative baseline models, while ablation studies substantiate the effectiveness of the theory-guided feature selection and feature learning components.

2. Related work

This section reviews deep learning-based recommender systems and related research on web browsing behavior for personalized recommendation. It then introduces IFT as a framework for understanding and supporting information-intensive decision-making.

2.1. Web browsing patterns in personalized recommendation

Web browsing patterns serve as an informative feedback source for personalized recommendation systems, reflecting user preferences, content satisfaction, and decision-making processes during interactions. This subsection focuses on two widely studied dimensions of browsing patterns: page visit logs and dwell time.

Page visit logs, which capture users' actions and product characteristics, are widely used to model browsing patterns in recommendation. Liang et al. [21] propose a semantic-expansion approach that constructs user interest profiles from previously browsed documents. Wang et al. [22] model website navigation sessions as directed graphs, where nodes represent web pages and edges denote user click paths. Kim et al. [23] transform session-based clickstream data into product- and page-level graph structures, extracting graph metrics to capture complex browsing patterns. In contrast, dwell time complements page visit logs by distinguishing meaningful interactions from accidental clicks, thereby reflecting user engagement. Yi et al. [24] incorporate dwell time into learning-to-rank and collaborative filtering frameworks, demonstrating improvements over click-only signals. Zhao et al. [25] propose a duration-wise Gaussian Mixture Model to disentangle true user interest from biased watch time. TE-GNN [26] integrates dwell time and inter-visit intervals into session graphs to capture interest drift. Bogina and Kuflik [27] incorporate dwell time into recurrent neural networks for session-based recommendation, while Mokryn et al. [28] combine temporal signals and product popularity trends to infer purchase intention.

Despite some achievements, existing research still suffers from two limitations. First, existing research lack a principled theoretical lens to determine which information from page visit logs should be incorporated into preference learning. On one hand, some studies indiscriminately inject all available contextual information from page visit logs into recommender systems in pursuit of information completeness. This naive approach not only introduces massive redundant features and noise, but also incurs prohibitive computational overhead that hinders real-world deployment. On the other hand, popular sequential recommender systems only consider product ID as the sole input, discarding rich contextual signals contained within page visit logs. This oversimplification fails to model users' decision-making processes during their product-seeking journeys, resulting in biased user preferences. Second, existing research treat dwell time as a static weight for edges in session graphs and a positional bias term in sequence modeling, failing to explore its latent relationships with other contextual information in page visit logs. For example, accounting for behavior type and product category enables more fine-grained user preference learning.

2.2. Deep learning-based product recommendation

Deep learning has become the dominant paradigm for modeling complex nonlinear relationships in user–product interactions [29,30]. Recent studies propose increasingly sophisticated frameworks based on architectures such as Transformers [29,31] and graph neural networks [30,32]. However, data sparsity in user–product interactions continues to constrain performance. To address this issue, recent research leverages side information to enrich user and product representations, giving rise to subfields including sequential, temporal, multi-behavior, attribute-aware, and social recommendation. Specifically, sequential recommendation captures the evolution of user interests from historical sequences [2], while temporal recommendation incorporates dynamics such as inter-event intervals and preference periodicity [5]. Multi-behavior recommendation leverages diverse user actions, such as viewing and purchasing, to construct richer preference representations [3]. Attribute-aware recommendation integrates side information from products and interaction contexts, such as category and device, to improve performance [4]. Social recommendation exploits user relationships to infer preferences, grounded in theories such as homophily and social influence [6]. Despite substantial progress, many studies rely on a single feature-learning paradigm, resulting in partial feature utilization. Sequential models cannot capture affinitive relationships that naturally form graph structures, while graph-based models fail to represent the sequential patterns and temporal dynamics underlying user behavior. This limitation motivates hybrid approaches that integrate both paradigms effectively.

2.3. Information foraging theory for decision support

IFT provides a principled framework for understanding how humans search for information in complex environments [33]. Drawing on an analogy with optimal foraging theory in ecology, it conceptualizes humans as informavores who seek to maximize their rate of acquiring valuable information per unit cost [20]. This perspective has proven valuable across diverse decision support domains by offering predictive models of user behavior and guiding the design of more effective information systems [17–19]. Proximal cues and foraging costs are the two foundational constructs of IFT. Specifically, proximal cues provide value signals for users to assess potential patches' expected information gain, while foraging costs represent the cognitive resource consumption users incur during information seeking journey.

The application of IFT spans empirical research [34] and design science [35]. Empirical studies show that key information-seeking behaviors, such as link selection [36] and dwell time [37], can be modeled as cost–benefit optimizations governed by perceived information scent. In design science, IFT has informed the development of intuitive interfaces for software debugging tools [17] and image retrieval systems [38], where cues are designed to enhance information scent and reduce foraging costs.

In the context of product recommendation, IFT provides a lens for analyzing user decision-making beyond static preference matching. Users navigate product spaces by following information scent derived from proximal cues and perceived foraging costs, with products functioning as information patches. This scent reflects users' evaluation of the value and cost associated with different information sources. Their objective is to efficiently gather sufficient product information to decide which products to engage with, through which behaviors (e.g., viewing, adding to cart, purchasing), and for how long. Guided by this perspective, the three core modules of Tiresias are designed under the IFT framework. First, the IFT-guided feature selection module is based on the premise that user decisions are shaped by both cues and costs, rather than a limited subset of product attributes. Existing systems often emphasize product-side features such as category [39] and brand [40], which capture only partial value signals while overlooking peer-derived cues and foraging costs. In contrast, our approach systematically selects features that map to IFT constructs, ensuring that the input space reflects key decision drivers. Second, the IFT-guided user preference extractor is motivated by the inherently sequential nature of product foraging. Unlike static models that assume fixed preferences, IFT suggests that users' cost–benefit evaluations evolve throughout the foraging process. Accordingly, we design a sequential learning module that captures the dynamic evolution of user preferences across interactions. Finally, the IFT-guided multi-task learning module reflects the view that a user's final interaction is the outcome of a continuous cost–benefit trade-off. The primary next-product prediction task models this outcome, while auxiliary cue and cost prediction tasks encourage the model to capture the underlying decision mechanisms rather than relying solely on correlational patterns.

3. Artifact description

As a fundamental problem-solving paradigm, design science research focuses on creating and evaluating innovative artifacts to solve real-world practical problems [41]. Following the well-established criterion of design science research in recent works [42–44], we adopt the six-step DSRM framework of Peffers et al. [15] to develop our artifact Tiresias. This section presents a comprehensive description of our artifact design.

3.1. Problem statement

Given a user's sequential interaction history, the objective of product recommendation is to generate a ranked list of top-K products that the user is most likely to interact with next. Formally, let $U = \{u_1, \dots, u_{|U|}\}$ and $V = \{v_1, \dots, v_{|V|}\}$ denote the sets of users and products, respectively. For a user u_i, the interaction sequence is represented as $s_i = \{v^i_1, \dots, v^i_l\}$ of length l. Users and products are projected into a d-dimensional latent space, yielding initial embedding matrices $E_u \in \mathbb{R}^{|U|\times d}$ for users and $E_v \in \mathbb{R}^{|V|\times d}$ for products. In addition, contextual information F is incorporated to alleviate data sparsity and facilitate a deeper understanding of user intent. Therefore, the objective is to predict the next product v^i_{l+1} that best aligns with the preference of user u_i, conditioned on the interaction sequence s_i and contextual information F. To this end, the recommender assigns a relevance score to each candidate product and recommends the top-K ranked products.

3.2. The overall framework

As illustrated in Fig. 2, the proposed Tiresias framework comprises three key modules. First, the IFT-guided feature selection module partitions side information into proximal cues and foraging costs and parametrizes these information as learnable feature representations. Specially, the Node2Vec is applied to the user affinity network, capturing peer-derived foraging decision cues from users who share highly similar product foraging experiences. Second, the IFT-guided user preference extractor is designed to model the evolution of user preferences from their product-foraging journey. Serving as the core component of this extractor, the Typhon module employs a Kolmogorov–Arnold network-enhanced differential transformer with a noninvasive feature injection mechanism to iteratively update product representations. Subsequently, the unsqueeze and excitation network integrates short-term and long-term user intents with static user features to obtain a comprehensive contextual preference representation. Third, the prediction and IFT-guided optimization module refines user–product matching through a multi-task learning objective that combines a primary next-product prediction task with auxiliary tasks for predicting proximal cues and foraging costs. The technical details of each module are elaborated in the following subsections.

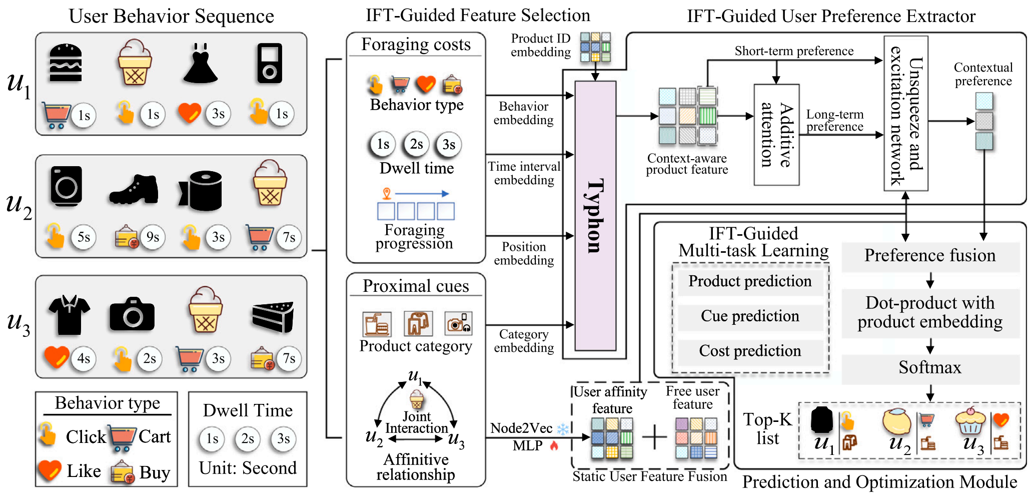


Fig. 2. Overview of the proposed Tiresias framework.

3.3. Information foraging theory-guided feature selection

Because cognitive resources for processing information are limited [45], users in a given context evaluate potential benefits relative to perceived costs by following the information scent of proximal cues. In this study, five contextual features are considered: behavior type, dwell time, foraging progression, product category, and affinitive relationship. To remain consistent with IFT, these features are organized into two conceptual categories: foraging costs and proximal cues. Specifically, we propose behavior type, dwell time, and foraging progression as foraging costs to represent users' resource consumption during their product foraging journey. Moreover, we propose product category and affinitive relationship as proximal cues to shape users' expected benefits of foraging behavior. The empirical support and operationalization for each feature are described below.

(i) **Behavior type as coarse-grained explicit foraging cost**. In the context of product recommendation, each distinct user behavior type corresponds to a heterogeneous level of cognitive effort and psychological risk, reflecting coarse-grained differences in the intensity of user interest [32]. For example, viewing or browsing behaviors demand minimal cognitive effort, as they only require users to scan superficial product information without deep evaluation or deliberate decision-making. These low-effort behaviors typically signal tentative, exploratory foraging, where users incur minimal costs to gather preliminary information about potential product patches [46]. In contrast, more involved behaviors, such as adding products to shopping carts [47] and saving products to wish lists [48], require substantially higher cognitive effort. Users need to assess the potential value of the product to evaluate whether it fits with the requirements. This evaluation process incurs non-trivial cognitive costs [32]. Moreover, each behavior type corresponds to a distinct level of risk exposure and associated psychological cost. For example, purchasing behavior requires a substantial cognitive cost [49], as it involves committing financial resources [50], confronting the perceived pain of paying [51], and bearing the psychological risk of post-purchase regret [52]. Additionally, most online retailers enforce explicit limits on virtual shopping cart capacity [53]. Every cart-addition action thus carries an inherent future foraging cost: reaching the capacity limit forces users to complete unplanned purchase or removal decisions to continue foraging [53]. This forced decision-making induces cognitive dissonance [54], whose resolution requires extra cognitive effort. Accordingly, by operationalizing behavior type as a foraging cost, our framework incorporates the cost-benefit trade-offs that govern user decision-making. Formally, let $B = \{b_1, \dots, b_{|B|}\}$ denote the set of all possible behaviors in a dataset, and let $E_b \in \mathbb{R}^{|B| \times d}$ be the corresponding behavior embedding matrix.

(ii) **Dwell time as fine-grained explicit foraging cost**. Unlike the coarse-grained cost differentiation captured by behavior type, dwell time provides fine-grained quantification of cognitive foraging costs across user-product interactions [55]. For online consumers, time saving is a primary motivation for choosing e-commerce channels [56]. For online information foragers, time is an inherently scarce resource. The time a user allocates to examining a specific product cannot be reallocated to exploring alternative product information patches, which generates an unavoidable opportunity cost of foraging [57]. Additionally, dwell time serves as an objective and widely validated surrogate measure of the cognitive effort dimension of foraging costs. For example, a negligible dwell time typically reflects minimal cognitive engagement (e.g., accidental clicks or superficial scanning), incurring lower cognitive foraging costs. Moreover, a sustained dwell time indicates deliberate and effortful product evaluation that imposes substantial cognitive costs on the user [57]. To effectively model dwell time, a time-interval encoding mechanism is introduced. Specifically, the time interval between two consecutive interactions is computed and then clipped at a predefined threshold $t_{\max}$ to mitigate the influence of outliers caused by prolonged inactivity or background usage. Formally, $\Delta_{i,j} = \min(t_{\max}, |t_j - t_i|)$, where t_i and t_j denote the timestamps of interactions. They are then mapped into an embedding matrix $E_t \in \mathbb{R}^{(t_{\max}+1) \times d}$, where zero index serves as a padding token.

(iii) **Foraging progression as implicit foraging cost**. Unlike explicit costs (immediate, single-interaction effort), foraging progression serves as a proxy for implicit costs, capturing latent cognitive burdens (e.g., perceived time, effort, and complexity [58]) accumulated along the foraging trail [37]. As the journey progresses, cumulative evaluation load pushes users toward cognitive limits, increasing overload risk and reducing decision quality [59]. For example, with identical explicit costs (e.g., 10-second dwell time), evaluating the fifteenth product incurs higher implicit cost than the second, as prior evaluations deplete cognitive resources [60]. In addition, primacy and recency effects [61] lead to stronger recall and preference for early and late items. This variation is captured through foraging progression, typically operationalized via learnable positional encoding. Formally, a positional embedding matrix $E_p \in \mathbb{R}^{l \times d}$ is defined, where each embedding represents the latent cognitive burden of its position index.

(iv) **Product category as internal proximal cues**. Product category captures intrinsic characteristics of products and provides a foundational semantic description, which can be viewed as internal proximal cues. Specifically, it serves as the critical knowledge-organizing tools that fundamentally shape how consumers interpret, evaluate, and make sense of products [62]. Additionally, investigating users' attention to product categories offers a valuable lens through which to gain insights into the cognitive factors associated with their foraging behaviors. For example, related literature suggests that utilitarian browsing behavior is defined by a focused search within a specific product category, as users rely on category boundaries to narrow their foraging scope and avoid irrelevant information [63]. In contrast, hedonic browsing behavior, which is dominated by exploratory search, is significantly less focused, with users allocating more of their foraging session to viewing broader category-level pages rather than individual product pages [63]. This distinction in user behavior demonstrates that product category acts as a proximal cue that signals the scope and relevance of information patches, allowing users to select and adjust their foraging strategy before committing to costly product-level evaluation. To operationalize the internal proximal cues, we adopt learnable category encoding. Formally, let $C = \{c_1, \dots, c_{|C|}\}$ denote the set of categories. Each product $v_i \in V$ is associated with a category $c_k \in C$. These categories are mapped into a hidden feature matrix $E_c \in \mathbb{R}^{|C|\times d}$ to capture category-aware preferences.

(v) **Affinitive relationship as external proximal cues.** Compared to product category as the internal proximal cues, affinitive relationship is conceptualized as external proximal cues that signals the foraging paths of similar users to deliver valuable information scent. This operationalization is reflected in the pervasive design of modern e-commerce platforms, which feature co-browsing recommendation modules with headings such as "90% of users who viewed this product also viewed" or "similar users also viewed" to surface this peer-derived signal to foraging users. From the perspective of herd behavior theory, users routinely draw on the opinions and observed foraging behavior of others to inform their own purchase decisions [64]. It often appears when consumers lack confidence in their domain-specific product knowledge [65], or when they are unwilling to expend extensive cognitive effort on low-involvement products [66]. In fact, consumers often assume that other users have superior information regarding product quality and value, and therefore seek to align their own choices with the collective behavior of their peers in order to minimize decision risk [67]. Additionally, homophily theory posits that users tend to trust and are more likely influenced by similar others [68]. Accordingly, users often seek opinions from peers with similar experiences when making decisions, and rely less on their own private information [69]. To this end, affinitive relationships are extracted from user–product interactions by computing the Jaccard similarity between pairs of users [70]. Formally, $\mathrm{sim}_{ij} = \frac{|s_i \cap s_j|}{|s_i \cup s_j|}$, where s_i and s_j denote the interaction sequences of users u_i and u_j, respectively. To filter out low-confidence connections, only the top 1% reliable relationships are retained. After extracting affinitive relationships, we construct a static user affinity network $\mathcal{G}_{an} = (\mathcal{V}_u, \mathcal{E}_{hm})$, where $\mathcal{V}_u$ denotes the set of user nodes and $\mathcal{E}_{hm}$ represents the set of edges corresponding to high-confidence affinitive relationships. On top of this network, Node2Vec [71] is employed to derive user affinity features $E_{hm} \in \mathbb{R}^{|U|\times d}$ that encode the influence of similar users in the interaction space. To preserve these static relational priors, E_{hm} is frozen during subsequent model training. This parameter freezing not only reduces the effective model complexity and mitigates the risk of overfitting, but also enables the downstream sequential component to concentrate on capturing users' evolving dynamic preferences rather than relearning stable relational structure.

3.4. Information foraging theory-guided user preference extractor

Fig. 3 shows that Typhon builds on Transformer-based models [29, 31,43] and is tailored to model users' foraging patterns by integrating contextual features, including behavior type, dwell time, foraging progression, and product category. Rather than adopting a vanilla Transformer, it introduces components for heterogeneous signals. Central to this design is a noninvasive feature injection strategy that preserves product representations while enabling contextual guidance. Specifically, query and key are derived from a mixture of product and contextual features, whereas value is computed solely from product embeddings:

$$\bar{Q}_1 = (E_b^{seq} + E_t^{seq} + E_c^{seq} + E_v^{seq})W_Q^1, \tag{1}$$

$$\bar{K}_1 = (E_b^{seq} + E_t^{seq} + E_c^{seq} + E_v^{seq})W_K^1, \tag{2}$$

$$V_{ID} = E_v^{seq} W_V^1, \tag{3}$$

where $E_{(\cdot)}^{seq} \in \mathbb{R}^{l\times d}$ denotes the corresponding feature sequence retrieved from the initial embedding matrix according to the product IDs in the foraging trail, and $W_Q^1, W_K^1, W_V^1 \in \mathbb{R}^{d\times d}$ are trainable parameters. Motivated by recent advances in large language models (LLMs), we employ rotary position embeddings (RoPE) [72] instead of standard learnable positional encodings, owing to their superior training stability and generalization. In particular, RoPE applies a position-dependent rotation matrix to the query and key vectors, enabling the model to capture absolute and relative positional dependencies simultaneously.

RoPE also offers an important context-aware advantage. Standard learnable positional encodings are detached from the information foraging context, whereas RoPE is applied within the self-attention mechanism directly to the context-enriched query and key vectors. Consequently, the rotation introduced by RoPE encodes positional information in a manner that is conditioned on the user's current foraging state. For a query vector $\bar{q}_m$ at position m and a key vector $\bar{k}_n$ at position n, RoPE rotates them via rotary matrices M_m and M_n, respectively, as follows:

$$RoPE(\bar{q}_m)^T \cdot RoPE(\bar{k}_n) = (M_m\bar{q}_m)^T(M_n\bar{k}_n) = \bar{q}_m^T M_{n-m}\bar{k}_n, \tag{4}$$

$$M_m = \begin{bmatrix} M_{m\theta_1} & & \\ & \ddots & \\ & & M_{m\theta_{d/2}} \end{bmatrix}, \quad M_{m\theta_i} = \begin{bmatrix} \cos m\theta_i & -\sin m\theta_i \\ \sin m\theta_i & \cos m\theta_i \end{bmatrix}, \tag{5}$$

where the selection of rotation angles satisfies the frequency $\theta_i = 10\,000^{-2(i-1)/d}$, $i \in [1, 2, \dots, d/2]$.

Subsequently, the attention distribution is computed via:

$$Attn_1 = \mathrm{Softmax}\left(\frac{Q_1 K_1^\top}{\sqrt{d}}\right) = \mathrm{Softmax}\left(\frac{\mathrm{RoPE}(\bar{Q}_1)\,\mathrm{RoPE}(\bar{K}_1)^\top}{\sqrt{d}}\right). \tag{6}$$

where $Attn_1 \in \mathbb{R}^{l\times l}$ denotes the learned primary attention distribution. Subsequently, we introduce differential attention mechanism to learn robust spiky attention distribution by calculating the difference between two separate softmax attention distribution [73].

To this end, we learn an additional background attention distribution $Attn_2 \in \mathbb{R}^{l\times l}$ in a similar manner. The refined attention distribution $Attn_{diff} \in \mathbb{R}^{l\times l}$ is then obtained as the element-wise difference:

$$Attn_{diff} = Attn_1 - \lambda Attn_2, \tag{7}$$

$$\lambda = \exp(\lambda_{Q1} \cdot \lambda_{K1}) - \exp(\lambda_{Q2} \cdot \lambda_{K2}) + \lambda_{init}, \tag{8}$$

$$\lambda_{init} = 0.8 - 0.6 \times \exp(-0.3 \times (f - 1)), \tag{9}$$

where λ controls differential adjustment strength, $\lambda_{Q1}, \lambda_{K1}, \lambda_{Q2}, \lambda_{K2} \in \mathbb{R}^d$ are learnable vectors, λ_{init} initializes λ, and f denotes the Typhon layer index. We follow the DIFF initialization strategy [74]. The differential operation balances the spectral distribution of attention matrices and mitigates rank collapse [74]. It also induces dynamic sparsity, encouraging attention to focus on more attractive products. The resulting $Attn_{diff}$ reflects the user's assessment of value and cost from proximal cues, encoding the information scent of each product. Finally, the representation over previously foraged products is $\tilde{E}v^{seq} = Attndiff V_{ID}$.

To enhance the nonlinear capacity of the differential self-attention network, GR-KAN [75] is incorporated. Based on the Kolmogorov–Arnold theorem, KAN [76] replaces fixed MLP activations with spline-based learnable functions on edges, enabling more expressive approximation. GR-KAN extends KAN with group-wise rational activations, improved GPU efficiency, parameter sharing, and variance-preserving

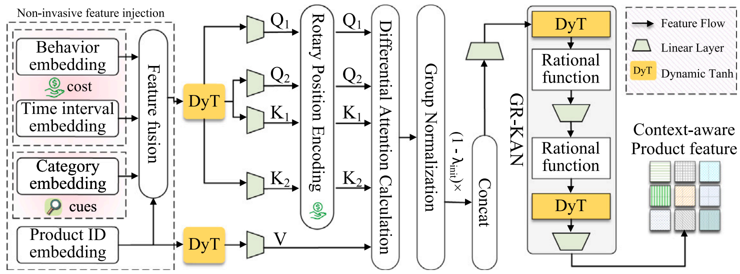


Fig. 3. Schematic view of our Typhon.

initialization for stable optimization. Formally, the GR-KAN module in Typhon is defined as

$$\hat{E}_v^{seq} = \left(\text{GR-KAN}(\text{GR-KAN}(\tilde{E}_v^{seq})W_2)\right)W_3, \tag{10}$$

$$\text{GR-KAN}(x) = \left[\sum_{i=1}^{d} \omega_{i,1} F_{\lfloor i/d_g \rfloor}(x_i) \quad \dots \quad \sum_{i=1}^{d} \omega_{i,d} F_{\lfloor i/d_g \rfloor}(x_i)\right], \tag{11}$$

$$F(x_i) = \frac{a_0 + a_1 x_i + \dots + a_m x_i^m}{1 + |b_1 x_i + \dots + b_n x_i^n|}, \tag{12}$$

where $W_2, W_3 \in \mathbb{R}^{d\times d}$ are trainable weight matrices, x_i denotes the ith input dimension, and g is the number of groups. Each group has $d_g = d_{\text{in}}/g$ dimensions, and the group index of x_i is given by $\lfloor i/d_g \rfloor$. The function $F(\cdot)$ is an input-wise rational activation based on the Safe Padé Activation Unit [77], with coefficients a_n and b_m selected to ensure stable training.

To improve both efficiency and stability, we also adopt dual dynamic Tanh [78,79]. Specifically, a dynamic Tanh (DyT) transformation is applied before the computation of the query, key, and value. For example, Eq. (3) can be rewritten as

$$V_{ID} = \text{DyT}(E_v^{seq})W_V^1, \tag{13}$$

$$\text{DyT}(x) = \gamma \cdot \text{Tanh}(\alpha x) + \beta, \tag{14}$$

where α is a learnable scalar parameter that allows scaling the input differently based on its range. The learnable and per-channel vector parameters β and γ allow the output to scale back to any scales. Moreover, we adopt DyT after stale differential attention calculation: $\tilde{E}_v^{seq} = \text{DyT}(Attn_{diff} V_{ID})$. Furthermore, we adopt dual dynamic Tanh in non-linear modeling component:

$$\hat{E}_v^{seq} = \text{DyT}\left(\text{GR-KAN}\left(\text{GR-KAN}(\text{DyT}(\tilde{E}_v^{seq}))W_2\right)\right)W_3. \tag{15}$$

For simplicity, the dropout operation, which is used to mitigate overfitting, is omitted from the equations.

Once context-aware product representations are obtained, a dual soft-attention network constructs user preference representations by jointly modeling short-term intent, long-term intent, and user-related features. Short-term intent is represented by the last product in the foraging sequence, i.e., $h_{short} = \hat{e}_{v_l}^{seq}$. Long-term intent is computed through an attention-weighted aggregation of the interaction sequence:

$$h_{long} = \sum_{v_i \in s_i} \delta_i \, \hat{e}_{v_i}^{seq}, \tag{16}$$

$$\delta_i = \text{Softmax}\left(W_\delta^\top \sigma\left(\left[W_4 \hat{e}_{v_i}^{seq} \,\|\, W_5(e_i^u + e_i^{hm}) \,\|\, W_6 h_{short}\right]\right)\right), \tag{17}$$

$$e_i^{hm} = MLP(\hat{e}_i^{hm}) \tag{18}$$

where $e_i^u \in E_u$ encodes the free (implicit) user preference and $\hat{e}_i^{hm} \in E_{hm}$ encodes the user affinity derived from affinitive relationships. $MLP(\cdot)$ denotes a three-layer MLP to refine the frozen user affinity feature, aligning with the contextual user preference during product foraging process. $\sigma(\cdot)$ denotes the sigmoid activation function. The matrices $W_4, W_5, W_6 \in \mathbb{R}^{d\times d}$ and $W_\delta \in \mathbb{R}^{3d}$ are learnable parameters of the attention network.

Subsequently, a second soft-attention module is used to fuse the three components into a contextual user preference representation h:

$$H_{fea} = \text{Stack}\left(W_7 h_{short}, \; W_8 h_{long}, \; W_9(e_i^u + e_i^{hm})\right), \tag{19}$$

$$\epsilon = \text{Softmax}(H_{fea}), \tag{20}$$

$$h = \text{SqueezeSum}(\epsilon \odot H_{fea}), \tag{21}$$

where $\odot$ indicates element-wise multiplication and $W_7, W_8, W_9 \in \mathbb{R}^{d\times d}$ are learnable weight matrices. Here, Stack$(\cdot)$ creates an extra channel dimension used to store the three transformed feature vectors. The attention weights ϵ are computed along this channel to quantify the importance of each component, and SqueezeSum$(\cdot)$ performs an attention-weighted summation over this channel followed by removing it, resulting in the final contextual preference vector h.

3.5. Prediction and IFT-guided optimization module

To predict the next product, this module calculate the interaction probability for each candidate product. Formally, $\hat{y}_i = \text{Softmax}(\text{sim}(h, e_{v_i})/\tau)$, where τ denotes the temperature parameter and $\text{sim}(\cdot)$ denotes the cosine similarity. $e_i^u \in E_v$ denotes the embedding feature of the candidate product v_i. Furthermore, we design a multi-task learning object function to optimize our Tiresias. The main task focuses on predicting the next product interaction, while the auxiliary tasks explicitly guide the model to develop a more nuanced understanding of the decision-making context. Formally,

$$\mathcal{L} = \mathcal{L}_{main} + \zeta(\mathcal{L}_{cue} + \mathcal{L}_{cost}), \tag{22}$$

$$\mathcal{L}_{main} = -log\left\{\frac{\exp\left(\text{sim}\left(h, e_{l+1}^+\right)/\tau\right)}{\sum_{v_j\in|V|}\exp\left(\text{sim}\left(h, e_{v_j}\right)/\tau\right)}\right\}, \tag{23}$$

$$\mathcal{L}_{cue} = -log\left\{\frac{\exp\left(\text{sim}\left(h, e_c^+\right)/\tau\right)}{\sum_{c_j\in|C|}\exp\left(\text{sim}\left(h, e_{c_j}\right)/\tau\right)}\right\}, \tag{24}$$

$$\mathcal{L}_{cost} = -log\left\{\frac{\exp\left(\text{sim}\left(h, e_b^+\right)/\tau\right)}{\sum_{b_j\in|B|}\exp\left(\text{sim}\left(h, e_{b_j}\right)/\tau\right)}\right\}, \tag{25}$$

where ζ is the loss weight for the auxiliary cue and cost prediction tasks. $e_{l+1}^+ \in E_v$ denotes the ground-truth product feature, while $e_c^+ \in E_c$ and $e_b^+ \in E_b$ represent the category and behavior features of the ground-truth product, respectively. To mitigate overfitting, we apply dropout to candidate product features. In addition, the normalized temperature-scaled cross-entropy loss improves feature alignment and uniformity, enhancing discrimination. This multi-task learning objective encourages latent representations to capture behavioral patterns and the cognitive processes of cue evaluation and cost assessment underlying information foraging.

3.6. Model design analysis

This section first provides a time complexity analysis of the current implementation of our model to evaluate its scalability and deployment feasibility. Then, we disentangle theory-driven design logic from implementation-specific choices to define the generalizable contribution of our work.

Three key theory-driven design logics form the irreducible theoretical core of our work. The first is systematic IFT-based feature categorization, which facilitates theory-driven feature discovery and classification, moving beyond the ad-hoc heuristic feature engineering common in existing recommendation systems. The second is the multi-task learning objective that aligns model training directly with IFT's core cost-benefit trade-off insight. The third is the explicit dynamic sequential modeling of user preference evolution, which captures the inherently temporal and progressive nature of product foraging journeys. Beyond these design logics, several components are included for superior performance and can be replaced in other implementations. Specifically, our Typhon, the KAN-enhanced differential transformer, improves the accuracy and efficiency of sequential preference modeling compared to vanilla transformer and recurrent neural networks. Moreover, the Node2Vec feature extraction method used for graph embedding can also be replaced with alternative graph neural networks without affecting the overall theoretical framework.

The time complexity of Tiresias comprises five components. The IFT-guided feature selection module has complexity $O(|U|\eta\rho(d+1))$, where η and ρ denote the number and length of random walks. Typhon has complexity $O(l_{max}^2 d + l_{max} d^2)$, where l_{max} is the maximum foraging trail length. The contextual preference extractor requires $O(l_{max} d + l_{max} d^2)$, and the prediction and IFT-guided optimization module requires $O(|V|d)$. Overall, the complexity is $O(|U|\eta\rho + (|V| + l_{max} + |U|\eta\rho + l_{max}^2)d + 2l_{max}d^2)$, where $|U|$ and $|V|$ dominate. In practice, pre-training is used to capture user affinity features, which are then frozen to reduce overhead during main training. In addition, attention acceleration methods (e.g., flash attention) and modern hardware (e.g., GPUs) further improve efficiency, enabling scalable deployment.

4. Experiments

To demonstrate the effectiveness of Tiresias, the evaluation follows the design science validity framework of Larsen et al. [16], which defines three knowledge claims: criterion, causal, and context. Criterion claims are validated through overall performance comparisons (Section 4.3), demonstrating superiority over state-of-the-art methods. Causal claims are examined via ablation studies (Section 4.4), isolating the contribution of each component. Context claims are established through robustness analysis (Section 4.6) and a foraging trail length-based user group study (Section 4.5), identifying conditions where the framework performs effectively. Finally, post-hoc analysis provides an intuitive understanding of the framework's operation.

4.1. Datasets, metrics and baseline models

Two publicly available e-commerce datasets from Kaggle are used: Cosmetics and Electronics Store. The Cosmetics dataset contains user event data from an online cosmetics shop, organized into five monthly subsets (Oct 2019–Feb 2020), with four interaction types: view, add-to-cart, remove-from-cart, and transaction.[2] Due to its large scale and inconsistent user and product identifiers across months, each subset is used independently. The Electronics Store dataset records sequential user interactions over approximately five months, including three behaviors: view, cart, and purchase.[3] Key statistics are summarized in Table 1.

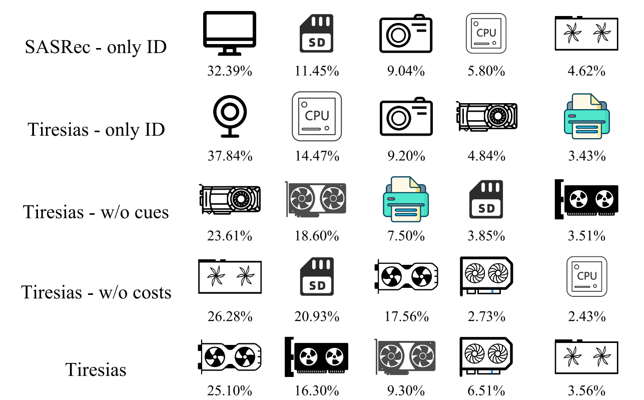


Fig. 4. Top-5 recommendation lists and predicted interaction probabilities of our Tiresias and four variants.

Recommendation performance is evaluated using Recall, MRR, and NDCG at cutoffs 5 and 10. To ensure unbiased evaluation, a full-ranking strategy is adopted [80], ranking the ground-truth item for each user against all others in the catalog. Interactions are ordered chronologically with a leave-one-out split: the last interaction for testing, the second last for validation, and the remainder for training. All experiments are repeated five times with different random seeds, and average results are reported.

Tiresias is benchmarked against 30 representative baselines across diverse recommendation domains (Table 2). The implementation builds on RecBole [81] and RecBole-GNN [82] to ensure reproducibility. Unless specified, the embedding dimension is 64, dropout 0.2, and temperature 0.07. The number of layers is tuned from 1, 2, 3, 4, 5, attention heads from 1, 2, 4, 8, loss weight from 1e–4, 1e–3, 1e–2, 1e–1, 1, time-interval threshold from 8, 16, 32, 64, 128, and foraging trail length from 3, 5, 10, 50, 100, 150. The AdamW optimizer [83] with a learning rate of 0.001 is used, with early stopping (patience 5) based on validation NDCG@10. For baselines, RecBole and SSLRec [84] and official code (when available) are used, reporting best performance; original hyperparameters are followed where possible and tuned otherwise.[4]

4.2. Visualized case study

This section presents a case study illustrating how key components of Tiresias affect the final recommendation list. Tiresias and four variants are implemented on the Electronics Store dataset: (i) SASRec-only ID, which removes all contextual information and uses vanilla self-attention; (ii) Tiresias–only ID, which removes contextual information; (iii) Tiresias–w/o cues, which removes proximal cues; and (iv) Tiresias–w/o costs, which removes foraging costs. A random user is selected, and the top-5 recommendations from each model, with predicted interaction probabilities, are compared. Fig. 4 highlights the difficulty of modeling user interests with limited information. SASRec-only ID is sensitive to incidental actions and fails to capture the user's interest in processors. Proximal cues and foraging costs better capture user intentions by aligning with cognitive processes. Tiresias-only ID fails to detect the user's strong preference for video cards, while Tiresias–w/o cues and Tiresias–w/o costs miss fine-grained preferences for specific video card types.

[2] www.kaggle.com/datasets/mkechinov/ecommerce-events-history-in-cosmetics-shop.

[3] www.kaggle.com/datasets/mkechinov/ecommerce-events-history-in-electronics-store.

[4] The implementation of Tiresias is available at: https://anonymous.4open.science/r/Tiresias-0541.

Table 1
Statistics of the datasets.

Dataset	# users	# products	# view	# add-to-cart	# remove-from-cart	# purchase	# relationships	# category	Avg. trail length
Cosmetics 2019 Oct	399 664	41 899	1 862 164	1 232 385	762 110	245 624	922 009	490	10.26
Cosmetics 2019 Nov	368 232	43 419	2 076 132	1 311 807	925 481	322 417	1 330 539	491	12.59
Cosmetics 2019 Dec	370 154	44 624	1 728 331	927 124	664 655	213 176	1 329 312	482	9.55
Cosmetics 2020 Jan	410 073	45 484	2 037 608	1 148 323	815 024	263 797	1 545 573	482	10.40
Cosmetics 2020 Feb	391 055	48 579	1 953 586	1 148 694	812 409	241 993	1 252 833	487	10.63
Electronics Store	490 399	53 453	793 089	54 032	–	37 343	414 117	718	1.8

Table 2
Summary of the baseline models and our Tiresias.

Model	Foraging costs			Proximal cues		Framework		Main approaches
	Behavior type	Dwell time	Foraging progression	Product category	Homo. relationship	Sequence-based	Graph-based	
Pop	✗	✗	✗	✗	✗	✗	✗	Heuristic method
Item-KNN [85]	✗	✗	✗	✗	✗	✗	✗	KNN
LightGCN [86]	✗	✗	✗	✗	✗	✗	✓	LightGCN
HNECL [87]	✗	✗	✗	✗	✗	✗	✓	LightGCN, CL
GRU4Rec [88]	✗	✗	✗	✗	✗	✓	✗	GRU
NARM [89]	✗	✗	✗	✗	✗	✓	✗	GRU, Attention
Muffin [90]	✗	✗	✗	✗	✗	✓	✗	FFT
DiffuRec [91]	✗	✗	✗	✗	✗	✓	✗	Diffusion model
SocialLGN [8]	✗	✗	✗	✗	✓	✗	✓	LightGCN
SERec [92]	✗	✗	✗	✗	✓	✗	✓	HAN
DGRec [93]	✗	✗	✗	✗	✓	✓	✓	LSTM, GAT
SMLP4Rec [94]	✗	✗	✗	✓	✗	✓	✗	MLP
BERT4Rec [95]	✗	✗	✓	✗	✗	✓	✗	Attention
SASRec [12]	✗	✗	✓	✗	✗	✓	✗	Attention
CL4SRec [96]	✗	✗	✓	✗	✗	✓	✗	Attention, CL
AC-TSR [97]	✗	✗	✓	✗	✗	✓	✗	Attention
SGNN-HN [98]	✗	✗	✓	✗	✗	✓	✓	GGNN, Attention
GSAU [99]	✗	✗	✓	✗	✗	✓	✓	LightGCN, Attention
MSSR [100]	✗	✗	✓	✓	✗	✓	✗	Attention, CL
ECL-SR [101]	✗	✗	✓	✓	✗	✓	✗	Attention, CL, AL
TiSASRec [102]	✗	✓	✓	✗	✗	✓	✗	Attention
SS4Rec [103]	✗	✓	✗	✗	✗	✓	✗	Mamba
KMCLR [104]	✓	✗	✗	✗	✗	✗	✓	Tatec [105], TransR [106], LightGCN
CML [107]	✓	✗	✗	✗	✗	✗	✓	LightGCN, CL
MBA [108]	✓	✗	✗	✗	✗	✗	✓	LightGCN
HAN-Full [13]	✓	✗	✗	✓	✓	✗	✓	HAN
MBHT [109]	✓	✗	✓	✗	✗	✓	✓	Linformer [110], HGCN
MGPT [111]	✓	✗	✓	✗	✗	✓	✓	GCN, LinRec [112], Atten-Mixer [113]
DIFSR [114]	✓	✗	✓	✓	✗	✓	✗	Attention
SASRec-Full [12]	✓	✓	✓	✓	✗	✓	✗	Attention
Tiresias (Ours)	✓	✓	✓	✓	✓	✓	✓	Node2Vec, GR-KAN [75], Differential attention [74]

Note: KNN denotes k-nearest neighbor classification. CL denotes contrastive learning. FFT denotes fast Fourier transform. HAN denotes heterogeneous attention network. GAT denotes graph attention network. GGNN denotes gated graph neural network. AL denotes adversarial learning. HGCN denotes Hypergraph convolutional network.

4.3. Recommendation performance

The results for top-5 and top-10 recommendations are reported in Tables 3 and 4, respectively. Several key observations emerge. First, Tiresias achieves superior performance on both datasets and across all evaluation metrics. It consistently attains the best scores, with improvements over the strongest baseline being statistically significant. On average, Tiresias yields gains of 1.91% in Recall, 2.83% in MRR, and 3.15% in NDCG across all datasets, indicating strong robustness and generalizability under different data distributions.

Second, most baselines fail to benefit consistently from additional side information and sometimes even degrade. For example, SocialGCN and SERec underperform LightGCN, while TiSASRec, DIFSR, and SASRec-Full fail to outperform SASRec. This behavior can be attributed to two factors: (i) feature selection is often driven by availability rather than principled criteria, leading to the omission of semantically important signals; and (ii) existing models are largely confined to either sequential or graph-based architectures, each with inherent limitations. Sequential models (e.g., SASRec, Muffin) struggle to capture latent user–product dependencies, whereas graph models (e.g., LightGCN, SocialGCN) have difficulty modeling fine-grained sequential patterns and temporal dynamics. Although hybrids such as DGRec and GSAU attempt to combine sequential and graph learning, their performance remains limited, partly due to over-reliance on complex graph modules and increased overfitting risk. These observations motivate the IFT-guided feature selection strategy and hybrid feature learning in Tiresias, where the Typhon sequence module serves as the core for most selected features and a lightweight graph module provides user affinity features.

Third, contrastive learning-based variants generally underperform their base models despite their reported promise. LightGCN outperforms HNECL, and SASRec remains stronger than CL4SRec and MSSR. This degradation is likely due to suboptimal negative sampling and data augmentations; for instance, the sequence-level augmentations used in CL4SRec can disrupt the semantic integrity of interaction sequences.

4.4. Ablation study

To further understand the contribution of each module, four groups of experiments are conducted on the Cosmetics 2019 Dec and Electronic

Table 3
Top-5 recommendation performance comparison of baseline models.

Models	Cosmetics 2019 Oct (@5)			Cosmetics 2019 Nov (@5)			Cosmetics 2019 Dec (@5)			Cosmetics 2020 Jan (@5)			Cosmetics 2020 Feb (@5)			Electronics store (@5)		
	Recall	MRR	NDCG	Recall	MRR	NDCG	Recall	MRR	NDCG	Recall	MRR	NDCG	Recall	MRR	NDCG	Recall	MRR	NDCG
Pop	0.55	0.25	0.32	1.94	1.44	1.56	2.61	2.06	2.19	2.29	1.78	1.91	1.72	1.22	1.34	4.36	2.59	3.02
Item-KNN	7.05	4.07	4.80	9.31	5.37	6.34	9.25	5.41	6.36	8.21	4.79	5.64	9.17	5.37	6.31	9.56	5.94	6.84
LightGCN	52.11	45.91	47.46	39.17	32.78	34.37	42.94	36.26	37.92	34.14	25.35	27.54	33.98	25.12	27.33	66.16	57.91	59.98
HNECL	50.25	42.46	44.41	31.62	24.22	26.07	37.84	29.38	31.49	35.28	27.00	29.07	35.86	27.65	29.70	64.77	50.85	54.33
GRU4Rec	58.67	50.22	52.33	43.70	32.67	35.42	46.62	35.13	38.00	45.88	35.04	37.75	46.12	34.98	37.76	71.60	62.02	64.42
NARM	58.76	50.42	52.50	43.63	32.68	35.41	46.50	35.43	38.19	45.61	35.18	37.78	45.99	35.18	37.88	72.35	61.85	64.48
Muffin	57.98	49.71	51.77	42.80	32.16	34.81	45.97	35.50	38.12	44.84	34.60	37.15	45.25	34.69	37.33	72.77	62.56	65.12
DiffuRec	55.82	47.54	49.60	40.57	29.76	32.46	44.06	32.87	35.67	41.04	31.12	33.59	42.70	32.18	34.81	69.83	59.86	62.36
SocialLGN	38.75	29.92	32.12	27.27	18.51	20.69	36.96	28.47	30.60	36.85	29.07	31.01	37.67	31.02	32.68	4.39	2.59	3.04
SERec	8.64	4.91	5.84	9.63	5.32	6.38	9.62	5.55	6.56	9.74	5.49	6.57	9.48	5.37	6.39	51.15	38.48	41.85
DGRec	11.73	6.12	7.53	10.22	5.55	6.69	11.32	6.27	7.49	10.99	6.76	7.23	10.11	5.41	6.57	59.55	36.05	42.03
SMLP4Rec	56.66	47.69	49.93	40.57	29.37	32.17	43.90	32.30	35.19	43.17	32.13	34.89	43.19	31.68	34.55	71.62	60.59	63.36
BERT4Rec	52.13	43.62	45.74	34.95	24.80	27.33	39.25	28.53	31.20	37.62	27.56	30.07	38.09	27.67	30.27	61.77	51.94	54.40
SASRec	57.94	50.40	52.28	44.05	33.22	35.92	47.33	36.36	39.10	46.23	35.89	38.47	46.26	35.71	38.35	72.40	62.46	64.94
CL4SRec	52.77	45.85	47.58	35.83	27.06	29.25	40.89	32.02	34.24	39.43	31.07	33.15	39.19	30.63	32.77	69.52	60.63	62.86
AC-TSR	58.14	50.49	52.39	43.11	33.09	35.59	46.15	36.04	38.56	45.21	35.58	37.98	45.62	35.57	38.08	72.03	62.42	64.82
SGNN-HN	58.76	49.86	52.08	43.58	31.84	34.77	46.74	34.65	37.67	45.75	34.29	37.15	45.93	34.28	37.19	72.99	61.86	64.65
GSAU	58.77	50.94	52.89	43.76	33.53	36.08	46.73	36.32	38.91	45.91	35.94	38.43	42.92	33.35	35.74	51.81	39.31	42.46
MSSR	58.05	49.02	51.28	42.78	32.25	34.88	45.16	34.46	37.13	44.98	34.68	37.25	44.97	34.38	37.03	72.05	62.06	64.56
ECL-SR	38.11	28.32	30.76	38.39	28.80	31.19	41.30	31.29	33.79	40.15	30.69	33.05	40.54	30.79	33.22	65.25	62.51	63.19
TiSASRec	58.23	50.59	52.49	42.90	32.98	35.45	46.29	36.06	38.61	45.32	35.68	38.08	45.60	35.55	38.05	72.04	62.42	64.82
SS4Rec	59.75	51.12	53.28	44.84	33.80	36.56	48.00	37.00	39.75	46.95	36.48	39.09	47.56	36.61	39.35	72.00	61.98	64.49
KMCLR	0.88	0.43	0.54	2.06	1.43	1.59	2.84	2.03	2.23	2.40	1.75	1.91	1.76	1.26	1.38	4.35	1.95	2.54
CML	1.61	0.80	1.00	2.85	1.48	1.82	3.36	2.22	2.50	3.00	1.65	1.99	2.26	1.10	1.39	4.35	2.46	2.92
MBA	6.10	3.52	4.15	6.25	3.74	4.36	4.62	3.00	3.40	4.14	2.65	3.02	3.60	2.39	2.69	2.71	1.79	2.01
HAN-Full	50.97	46.79	47.83	37.92	33.79	34.82	42.06	38.15	39.12	45.65	35.31	37.88	40.02	36.09	37.07	65.58	64.57	64.82
MBHT	24.01	15.06	17.28	33.10	25.34	27.27	22.82	14.73	16.74	34.41	26.97	28.82	20.23	12.90	14.72	8.03	4.84	5.62
MGPT	24.49	15.31	17.58	21.42	13.54	15.50	20.85	13.11	15.03	20.59	13.11	14.96	21.91	14.10	16.03	7.56	4.40	5.19
DIFSR	50.51	35.36	39.14	38.36	26.45	29.42	40.78	28.55	31.60	39.84	27.91	30.88	40.54	28.37	31.40	70.46	58.84	61.75
SASRec-Full	56.11	47.86	49.91	43.55	32.94	35.58	48.14	37.05	39.82	39.85	36.01	36.97	45.63	35.11	37.74	72.79	62.75	65.27
Tiresias (Ours)	**60.68*****	**52.46*****	**54.51*****	**46.50*****	**35.64*****	**38.35*****	**49.65*****	**38.66*****	**41.41*****	**48.05*****	**37.65*****	**40.25*****	**48.24*****	**37.54*****	**40.22*****	**74.34*****	**65.36*****	**67.61*****
Improvement	1.56%	2.62%	2.31%	3.70%	5.44%	4.90%	3.14%	1.34%	3.99%	2.34%	3.21%	2.97%	1.43%	2.54%	2.21%	1.85%	1.22%	3.59%

For each dataset, the **bold**-faced number is the best score and the second performer is underlined.

*** Indicates that our model significantly outperforms the best baseline model, based on a paired t-test with a p-value < 0.001.

Table 4
Top-10 recommendation performance comparison of baseline models.

Models	Cosmetics 2019 Oct (@10)			Cosmetics 2019 Nov (@10)			Cosmetics 2019 Dec (@10)			Cosmetics 2020 Jan (@10)			Cosmetics 2020 Feb (@10)			Electronics store (@10)		
	Recall	MRR	NDCG	Recall	MRR	NDCG	Recall	MRR	NDCG	Recall	MRR	NDCG	Recall	MRR	NDCG	Recall	MRR	NDCG
Pop	2.05	0.44	0.79	3.22	1.60	1.97	3.94	2.22	2.61	3.88	1.98	2.40	3.08	1.39	1.77	6.91	2.92	3.84
Item-KNN	10.34	4.50	5.86	13.72	5.96	7.77	13.56	5.99	7.75	12.21	5.32	6.92	13.30	5.91	7.64	13.23	6.42	8.02
LightGCN	56.14	46.45	48.76	43.91	33.41	35.90	47.63	36.88	39.44	39.82	26.11	29.38	39.65	25.88	29.16	69.86	58.41	61.18
HNECL	54.63	43.04	45.83	37.03	24.94	27.81	43.73	30.17	33.40	41.15	27.78	30.97	41.56	28.42	31.55	72.63	51.90	56.88
GRU4Rec	63.75	50.90	53.98	50.20	33.54	37.54	53.18	36.01	40.13	52.25	35.90	39.82	52.56	35.84	39.85	76.22	62.65	65.93
NARM	63.93	51.11	54.17	50.18	33.56	37.53	53.12	36.32	40.34	51.93	36.03	39.83	52.10	36.02	39.93	77.70	62.57	66.22
Muffin	63.12	50.40	53.44	49.31	33.04	36.93	52.67	36.40	40.29	51.23	35.46	39.22	51.75	35.56	39.43	78.04	63.27	66.83
DiffuRec	60.75	48.20	51.20	46.96	30.62	34.53	50.57	33.75	37.78	47.10	31.93	35.56	48.92	33.02	36.83	74.41	60.48	63.86
SocialLGN	45.27	30.79	34.23	33.81	19.37	22.80	42.23	29.18	32.30	41.73	29.72	32.59	42.19	31.63	34.15	6.82	2.91	3.81
SERec	12.19	5.36	6.92	13.80	5.87	7.73	13.41	6.01	7.68	13.35	5.99	7.81	13.20	5.88	7.58	54.02	38.73	42.54
DGRec	17.49	6.85	9.33	15.81	6.24	8.39	17.52	6.99	9.27	16.67	6.06	8.97	15.62	6.14	8.35	69.66	37.43	45.34
SMLP4Rec	62.11	48.42	51.7	47.61	30.32	34.45	50.97	33.24	37.48	49.99	33.05	37.10	50.20	32.62	36.82	77.06	61.32	65.12
BERT4Rec	57.30	44.31	47.42	41.62	25.70	29.49	46.30	29.48	33.49	44.10	28.45	32.23	44.81	28.57	32.44	66.79	52.62	56.03
SASRec	62.96	51.07	53.90	50.69	34.11	38.07	54.08	37.27	41.29	52.71	36.76	40.57	52.90	36.60	40.50	77.91	63.21	66.74
CL4SRec	56.97	46.41	48.94	41.40	27.81	31.05	46.24	32.74	35.97	44.69	31.78	34.86	44.55	31.36	34.51	74.11	61.25	64.35
AC-TSR	63.14	51.16	54.01	49.53	33.95	37.67	52.70	36.92	40.68	51.49	36.43	40.02	51.99	36.42	40.14	77.59	63.16	66.62
SGNN-HN	64.24	50.60	53.86	50.66	32.79	37.07	53.86	35.61	39.98	52.64	35.22	39.39	52.92	35.23	39.46	78.45	62.60	66.42
GSAU	63.86	51.63	54.54	50.29	34.41	38.20	53.33	37.20	41.05	52.30	36.80	40.50	49.05	34.17	37.73	56.82	39.99	44.09
MSSR	62.9	49.66	52.85	49.16	33.11	36.95	51.62	35.32	39.22	51.17	35.51	39.25	51.28	35.23	39.07	77.57	62.80	66.36
ECL-SR	44.69	29.20	32.89	44.87	29.67	33.29	47.71	32.15	35.86	46.41	31.53	35.08	46.97	31.65	35.31	66.84	62.72	63.71
TiSASRec	63.22	51.26	54.11	49.38	33.85	37.55	52.89	36.95	40.75	51.62	36.52	40.12	51.94	36.40	40.11	77.66	63.18	66.65
SS4Rec	64.98	51.82	54.97	51.44	34.69	38.70	54.63	37.89	41.90	53.50	37.36	41.22	54.08	37.49	41.46	77.25	62.69	66.19
KMCLR	0.88	0.49	0.68	3.09	1.56	1.92	3.73	2.14	2.51	3.33	1.87	2.20	3.00	1.41	1.77	6.72	2.26	3.31
CML	2.54	0.92	1.29	4.05	1.64	2.20	4.30	2.34	2.80	4.05	1.79	2.32	3.16	1.22	1.68	6.49	2.73	3.60
MBA	8.91	3.88	5.05	9.05	4.11	5.26	6.25	3.22	3.93	5.70	2.86	3.52	4.84	2.55	3.09	3.53	1.90	2.28
HAN-Full	53.83	47.17	48.76	41.06	34.21	35.83	45.00	38.54	40.07	52.28	36.20	40.03	42.91	36.48	38.00	66.57	64.70	65.14
MBHT	31.50	16.05	19.70	38.83	26.11	29.12	29.18	15.58	18.79	39.98	27.71	30.62	26.34	13.72	16.69	10.90	5.21	6.54
MGPT	32.10	16.32	20.04	27.68	14.38	17.52	27.12	13.94	17.05	26.75	13.93	16.96	27.94	14.90	17.98	10.71	4.81	6.20
DIFSR	58.53	36.44	41.75	45.66	27.43	31.79	48.44	29.58	34.08	47.24	28.90	33.29	47.93	29.37	33.81	75.89	59.58	63.52
SASRec-Full	61.22	48.54	51.57	50.45	33.86	37.82	54.54	37.91	41.90	42.86	36.41	37.94	52.50	36.03	39.96	78.05	63.46	66.98
Tiresias (Ours)	**65.68*****	**53.13*****	**56.14*****	**52.91*****	**36.50*****	**40.43*****	**56.04*****	**39.52*****	**43.48*****	**54.31*****	**38.49*****	**42.27*****	**54.53*****	**38.39*****	**42.26*****	**78.52*****	**65.92*****	**68.97*****
Improvement	1.08%	2.53%	2.13%	2.86%	5.22%	4.47%	2.58%	2.54%	3.77%	1.51%	3.02%	2.55%	0.83%	2.40%	1.93%	0.09%	1.89%	2.97%

For each dataset, the **bold**-faced number is the best score and the second performer is underlined.

*** Indicates that our model significantly outperforms the best baseline model, based on a paired t-test with a p-value < 0.001.

store datasets, examining the effectiveness of the feature selection approach, the feature learning method, the Typhon module, and the IFT-guided multi-task learning.

Effect of the feature selection approach. Three model variants are constructed to examine the effects of the selected proximal cues and foraging costs from an information foraging perspective, with each variant removing or replacing a specific feature. For example, *w/o BT* removes behavior type features relative to Tiresias, while *Learnable PE* replaces RoPE with vanilla positional encoding. The results are summarized in Table 5. The ablation results substantiate the necessity of each IFT-guided feature. Performance consistently degrades when any feature is removed or replaced, thereby validating the proposed feature selection approach. At the same time, the relative importance of specific features exhibits clear dataset dependence. On the *Cosmetics* dataset, which exhibits rich user interaction patterns, behavior type emerges as the most influential cognitive cost. On the *Electronics Store* dataset, RoPE yields greater gains than vanilla learnable positional

Table 5
The effect of feature selection approach.

Dataset	Model	Behavior type	Dwell time	RoPE	Product attribute	Affinitive relationships	@5			@10		
							Recall	MRR	NDCG	Recall	MRR	NDCG
Cosmetics 2019 Dec	w/o BT	✗	✓	✓	✓	✓	46.72	35.85	38.57	53.14	36.72	40.65
	w/o DT	✓	✗	✓	✓	✓	48.37	37.27	40.05	54.64	38.12	42.08
	Learnable PE	✓	✓	✗	✓	✓	47.91	37.06	39.77	54.41	37.94	41.88
	w/o PA	✓	✓	✓	✗	✓	48.78	37.82	40.56	55.16	38.68	42.63
	w/o HR	✓	✓	✓	✓	✗	48.20	37.34	40.05	54.67	38.21	42.15
	Tiresias (Ours)	✓	✓	✓	✓	✓	**49.65**	**38.66**	**41.41**	**56.04**	**39.52**	**43.48**
Electronics store	w/o BT	✗	✓	✓	✓	✓	71.81	62.32	64.70	76.25	62.93	66.15
	w/o DT	✓	✗	✓	✓	✓	72.12	62.57	64.97	76.63	63.18	66.43
	Learnable PE	✓	✓	✗	✓	✓	71.24	62.03	64.34	75.68	62.62	65.78
	w/o PA	✓	✓	✓	✗	✓	71.67	62.39	64.72	76.10	62.99	66.16
	w/o HR	✓	✓	✓	✓	✗	71.91	62.51	64.87	76.34	63.11	66.31
	Tiresias (Ours)	✓	✓	✓	✓	✓	**74.34**	**65.36**	**67.61**	**78.52**	**65.92**	**68.97**

For each dataset, the **bold**-faced number is the best score.

Table 6
The effect of overall framework.

Dataset	Model	Graph learning	Sequence learning	Preference fusion	@5 Recall	@5 MRR	@5 NDCG	@10 Recall	@10 MRR	@10 NDCG
Cosmetics 2019 Dec	w/o GL	✗	✓	✗	48.20	37.34	40.05	54.67	38.21	42.15
	GL-LightGCN	LightGCN	Typhon	Sum	48.93	37.82	40.60	55.50	38.70	42.72
	GL-GAT	GAT	Typhon	Sum	48.77	37.48	40.30	55.26	38.35	42.40
	GL-Cues	Metapath2Vec ❄	Typhon	Sum	46.97	35.95	38.69	54.05	36.90	40.99
	w/o SL	Node2Vec ❄	✗	✗	39.74	36.10	37.00	42.63	36.48	37.93
	SL-GRU	Node2Vec ❄	GRU	Sum	46.04	35.29	37.97	52.42	36.15	40.05
	SL-Mamba	Node2Vec ❄	Mamba	Sum	47.52	36.77	39.45	53.96	37.63	41.54
	Fusion-Attn	Node2Vec ❄	Typhon	Attention	47.86	37.01	39.72	54.27	37.87	41.80
	Fusion-MLP	Node2Vec ❄	Typhon	MLP	47.83	36.91	39.63	54.30	37.78	41.73
	Tiresias (Ours)	Node2Vec ❄	Typhon	Sum	**49.65**	**38.66**	**41.41**	**56.04**	**39.52**	**43.48**
Electronics store	w/o GL	✗	✓	✗	71.91	62.51	64.87	76.34	63.11	66.31
	GL-LightGCN	LightGCN	Typhon	Sum	72.54	62.84	65.28	76.80	63.42	66.67
	GL-GAT	GAT	Typhon	Sum	72.48	62.84	65.26	76.93	63.45	66.71
	GL-Cues	Metapath2Vec ❄	Typhon	Sum	73.77	64.99	67.20	77.80	65.54	68.51
	w/o SL	Node2Vec ❄	✗	✗	65.81	64.18	64.59	66.75	64.31	64.90
	SL-GRU	Node2Vec ❄	GRU	Sum	71.82	61.91	64.40	76.61	62.56	65.95
	SL-Mamba	Node2Vec ❄	Mamba	Sum	71.11	61.92	64.23	75.54	62.52	65.66
	Fusion-Attn	Node2Vec ❄	Typhon	Attention	70.95	61.65	63.98	75.63	62.28	65.50
	Fusion-MLP	Node2Vec ❄	Typhon	MLP	70.63	61.71	63.95	75.05	62.31	65.39
	Tiresias (Ours)	Node2Vec ❄	Typhon	Sum	**74.34**	**65.36**	**67.61**	**78.52**	**65.92**	**68.97**

For each dataset, the **bold**-faced number is the best score.

encoding. A plausible explanation is that RoPE rotates the context-enriched query and key vectors to distinguish implicit costs based on the user's current foraging state, whereas learnable positional encoding merely adds a static, context-agnostic bias and fails to capture how positional significance varies with the encountered context.

Effect of the feature learning method. In the second group of ablation studies, eight model variants are constructed to dissect the contribution of each feature learning module in Tiresias. Specifically, *w/o GL* removes the Node2Vec to capture user affinity feature, while *GL-LightGCN* and *GL-GAT* replace Node2Vec with LightGCN and GAT, respectively. *GL-Cues* adopts Metapath2Vec [115] on the heterogeneous graph to encode proximal cue features, while integrating all foraging costs into sequence learning to capture dynamic user preferences. Similarly, *w/o SL* removes the sequence learning module Typhon, while *SL-GRU* and *SL-Mamba* replace Typhon with GRU and Mamba [116], respectively. In addition, *Fusion-Attn* and *Fusion-MLP* adopt an attention network and an MLP, respectively, to fuse the free user feature and the user affinity feature for subsequent preference learning. The results are reported in Table 6, from which several insights emerge. First, the graph learning component is essential: removing it (*w/o GL*) leads to noticeable performance drops on both datasets. Replacing the lightweight Node2Vec with more expressive GNNs (LightGCN, GAT) fails to yield further gains, despite their stronger relational modeling capacity. This suggests that a simple frozen graph embedding is sufficient to capture stable user affinity features, while avoiding the overfitting risks associated with parameter-heavy graph networks. Second, the sequence learning module emerges as the most critical component. Its removal (*w/o SL*) substantially degrades performance, particularly on the *Cosmetics* dataset where dynamic preferences are more pronounced, indicating that the user's information foraging process fundamentally follows a sequential decision pattern. Although GRU and Mamba are competitive alternatives, both *SL-GRU* and *SL-Mamba* fail to match Typhon in modeling the trajectory of user preference evolution. Third, the comparison of preference fusion strategies yields an unexpected result. Despite their theoretical flexibility, both attention-based and MLP-based fusion (*Fusion-Attn*, *Fusion-MLP*) underperform the simple summation scheme. This finding suggests that these more complex fusion mechanisms tend to overfit, whereas the sum operation offers a more robust way to combine user features while preserving their distinct semantic contributions. Last, despite its theoretical coherence in aligning modeling mechanisms for features belonging to the same proximal cue category, *GL-Cues* yields a notable performance degradation relative to the main Tiresias model across all evaluation metrics and both datasets. This can be explained that while affinitive relationships and product-category attributes are both classified as proximal cues under the IFT framework, they exhibit fundamental heterogeneity in their information properties and functional roles. Specifically, user affinity relationships act as global and stable external proximal cues that signal validated foraging paths of similar peer users, while category attributes function as dynamic and context-dependent internal proximal cues that update incrementally with each step of a user's sequential foraging journey. Jointly modeling these two functionally distinct signal types in a static heterogeneous graph dilutes the unique information scent of each cue, leading to less discriminative latent representations.

Effect of the Typhon module. To investigate the Typhon's design, each of its core components is removed or replaced to construct four variants: *w/o DM* removes the differential mechanism; *w/o dual DyT* replaces DyT with vanilla layer normalization; *Invasive injection* directly sums all selected features to compute query, key, and value; and *NF-MLP* replaces GR-KAN with a standard MLP. As shown in Table 7, removing or substituting any component of Typhon leads to suboptimal performance on both datasets. The differential mechanism emerges as

Table 7
The effect of main components of Typhon.

Dataset	Model	Differential mechanism	Dual DyT	Non-invasive injection	Non-linear function	@5 Recall	@5 MRR	@5 NDCG	@10 Recall	@10 MRR	@10 NDCG
Cosmetics 2019 Dec	w/o DM	✗	✓	✓	GR-KAN	43.48	33.42	35.93	49.80	34.27	37.98
	w/o dual DyT	✓	✗	✓	GR-KAN	48.86	37.59	40.40	55.30	38.46	42.50
	Invasive injection	✓	✓	✗	GR-KAN	47.73	36.84	39.56	54.19	37.71	41.66
	NF-MLP	✓	✓	✓	MLP	48.51	37.41	40.18	55.05	38.29	42.31
	Tiresias (Ours)	✓	✓	✓	GR-KAN	**49.65**	**38.66**	**41.41**	**56.04**	**39.52**	**43.48**
Electronics store	w/o DM	✗	✓	✓	GR-KAN	70.77	61.54	63.86	75.04	62.12	65.25
	w/o dual DyT	✓	✗	✓	GR-KAN	71.98	62.45	64.84	76.64	63.08	66.36
	Invasive injection	✓	✓	✗	GR-KAN	71.37	62.05	64.38	76.07	62.68	65.91
	NF-MLP	✓	✓	✓	MLP	71.84	62.36	64.74	76.27	62.96	66.18
	Tiresias (Ours)	✓	✓	✓	GR-KAN	**74.34**	**65.36**	**67.61**	**78.52**	**65.92**	**68.97**

For each dataset, the **bold**-faced number is the best score.

Table 8
The effect of cue and cost prediction tasks.

Dataset	Model	Cue prediction	Cost prediction	@5 Recall	@5 MRR	@5 NDCG	@10 Recall	@10 MRR	@10 NDCG
Cosmetics 2019 Dec	w/o cue prediction	✗	✓	48.14	37.05	39.82	54.54	37.91	41.90
	w/o cost prediction	✓	✗	48.56	37.28	40.10	54.99	38.15	42.18
	w/o both aux tasks	✗	✗	48.59	37.66	40.39	55.04	38.52	42.48
	Tiresias (Ours)	✓	✓	**49.65**	**38.66**	**41.41**	**56.04**	**39.52**	**43.48**
Electronics store	w/o cue prediction	✗	✓	72.18	62.56	64.97	76.63	63.16	66.41
	w/o cost prediction	✓	✗	72.19	62.72	65.09	76.59	63.31	66.52
	w/o both aux tasks	✗	✗	72.14	62.59	64.98	76.65	63.20	66.45
	Tiresias (Ours)	✓	✓	**74.34**	**65.36**	**67.61**	**78.52**	**65.92**	**68.97**

For each dataset, the **bold**-faced number is the best score.

the most critical element, as its absence (*w/o DM*) results in the largest performance degradation. This can be attributed to the sparse attention distributions induced by the differential operation, which effectively assigns near-zero weights to irrelevant products within specific contexts. Dynamic Tanh is also indispensable: by implicitly emulating normalization, it applies near-linear transformations to small-valued features while nonlinearly squashing anomalous values. In addition, the noninvasive feature injection strategy clearly outperforms direct feature summation for query, key, and value computation, as it prevents rich contextual features from overwhelming product representations. Finally, GR-KAN consistently surpasses the vanilla MLP, underscoring its superior feature learning capacity and ability to preserve fine-grained variations.

Effect of the IFT-guided multi-task learning. Two groups of experiments are conducted to evaluate the effectiveness of the auxiliary tasks. First, an ablation study examines the individual and joint contributions of the cue and cost prediction tasks by comparing three variants: *w/o cue prediction*, which removes the category prediction task; *w/o cost prediction*, which removes the behavior prediction task; and *w/o both aux tasks*, which removes both cue and cost prediction. Second, the impact of different loss weights for the two auxiliary tasks is investigated. The ablation results in Table 8 show that both auxiliary tasks substantially improve overall performance by aligning preference learning with the user's cognitive process. The multi-task design is theoretically grounded in the user's cost–value trade-off during information foraging. The cue prediction task encourages the model to identify product categories, a key internal proximal cue used by users to estimate information value, while the cost prediction task focuses on forecasting behavior type, a direct manifestation of cognitive cost incurred during foraging. Together, these auxiliary tasks guide the model to internalize the underlying valuation and cost-assessment mechanisms that drive human decision-making.

The sensitivity analysis of the loss weight ζ, tuned in {1e–4, 1e–3, 1e–2, 1e–1, 1}, provides insights into balancing the auxiliary tasks. As shown in Fig. 5, both datasets achieve the best performance with a relatively small weight ($\zeta = 0.01$), suggesting that the auxiliary tasks should act as regularizers rather than dominate feature learning. When ζ is too large ($\zeta = 1$), performance degrades markedly — especially on the *Cosmetics* dataset — as the auxiliary objectives overwhelm the primary recommendation task. Conversely, extremely small values ($\zeta = 0.0001$) underutilize the benefits of the multi-task framework. This pattern is consistent with IFT: although cues and costs shape decision-making, the primary objective remains maximizing information value, i.e., accurate product prediction in this setting.

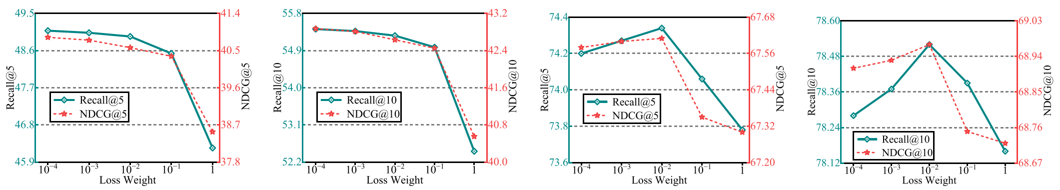
4.5. Effect of foraging trail length

To examine the impact of user engagement on recommendation performance, we analyze performance across different foraging trail lengths. Following recent work [117], the *Cosmetics 2019 Dec* dataset is divided into three groups: *Cold-start* (trail length ≤ 5), *Emerging* ($5 <$ trail length ≤ 10), and *Active* (trail length > 10). The statistics in Table 9 show a clear progression in user activity. Seven representative baselines are evaluated using Recall@10, MRR@10, and NDCG@10. As shown in Table 10, Tiresias consistently outperforms all baselines across the three groups, with statistically significant gains throughout. The largest improvements occur for *Active* users, with gains of 2.38% in Recall@10, 7.43% in MRR@10, and 5.76% in NDCG@10. The stronger gains in MRR and NDCG indicate that Tiresias more effectively ranks relevant products near the top for users with extensive interaction histories.

To determine the optimal foraging trail length for preference modeling, the maximum trail length is varied over 3, 5, 10, 20, 50, 100, 150. Fig. 6 shows dataset-specific optima: performance peaks at a trail length of 50 for *Cosmetics* and 20 for *Electronics Store*. This difference is consistent with the longer average trail length in *Cosmetics* (9.55) than in *Electronics Store* (1.8). Accordingly, Tiresias benefits from longer interaction histories to capture evolving preferences in *Cosmetics*, whereas recent interactions are sufficient for *Electronics Store*. Performance declines beyond the optimal trail length, likely because longer histories include outdated interactions that no longer reflect users' current preferences.

4.6. Robust analysis

A comprehensive robust analysis is conducted to identify suitable configurations for key hyperparameters and the user affinity network. In terms of key hyperparameters, the number of Typhon layers is searched over {1, 2, 3, 4, 5}, the number of attention heads over {1, 2, 4, 8}, the dwell time threshold over {8, 16, 32, 64, 128}, the batch size over {128, 256, 512, 1024, 2048}, and the embedding dimension over {8, 16, 32, 64, 128}. The results are shown in Fig. 7.



(a) Metrics@5 on Cosm. (b) Metrics@10 on Cosm. (c) Metrics@5 on Elec. (d) Metrics@10 on Elec.

Fig. 5. Tiresias performance with varying loss weights on the two datasets.

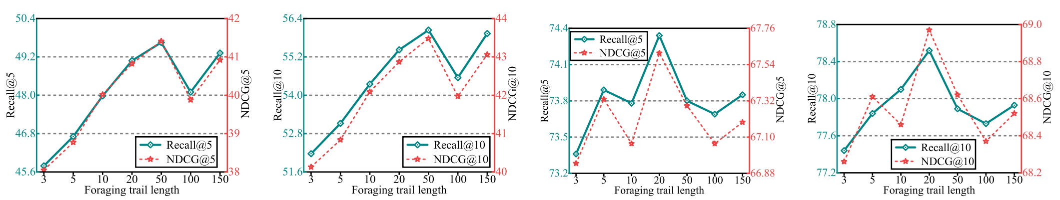
Table 9
Statistics of the three grouping datasets.

Dataset	# users	# products	# view	# add-to-cart	# remove-from-cart	# purchase	Avg. trail length
Cold-start group	289 868	30 517	445 976	35 863	4291	1879	1.68
Emerging group	26 417	23 605	136 201	46 829	11 363	5909	7.58
Active group	53 869	42 729	1 146 154	844 432	649 001	205 388	52.81

Table 10
Performance comparison for different user groups.

Model	Cold-start group (@10)			Emerging group (@10)			Active group (@10)		
	Recall	MRR	NDCG	Recall	MRR	NDCG	Recall	MRR	NDCG
SERec	47.69	37.50	40.77	29.88	22.37	24.18	15.51	6.00	8.21
SS4Rec	52.02	42.58	44.94	44.85	33.22	36.88	44.88	27.71	31.79
GSAU	34.21	24.84	27.10	34.95	16.70	21.01	40.68	26.79	30.09
MSSR	54.82	39.91	40.23	45.65	30.54	34.39	31.57	17.07	20.50
SGNN-HN	54.20	42.47	45.43	44.88	31.65	34.96	42.01	22.64	27.23
Muffin	54.75	39.05	42.83	46.83	31.03	34.81	41.28	26.04	29.66
DIFSR	51.45	32.50	37.05	40.24	23.95	27.83	32.60	16.80	20.53
Tiresias (Ours)	**55.84*****	**43.45*****	**46.48*****	**47.29*****	**33.88*****	**37.14*****	**45.95*****	**29.77*****	**33.62*****
Improvement	1.86%	2.04%	2.31%	0.98%	1.99%	0.70%	2.38%	7.43%	5.76%

For each dataset, the **bold**-faced number is the best score and the second performer is underlined.
*** Indicates that our model significantly outperforms the best baseline model, based on a paired t-test with a p-value < 0.001.



(a) Metrics@5 on Cosm. (b) Metrics@10 on Cosm. (c) Metrics@5 on Elec. (d) Metrics@10 on Elec.

Fig. 6. Tiresias performance with varying foraging trail length on the two datasets.

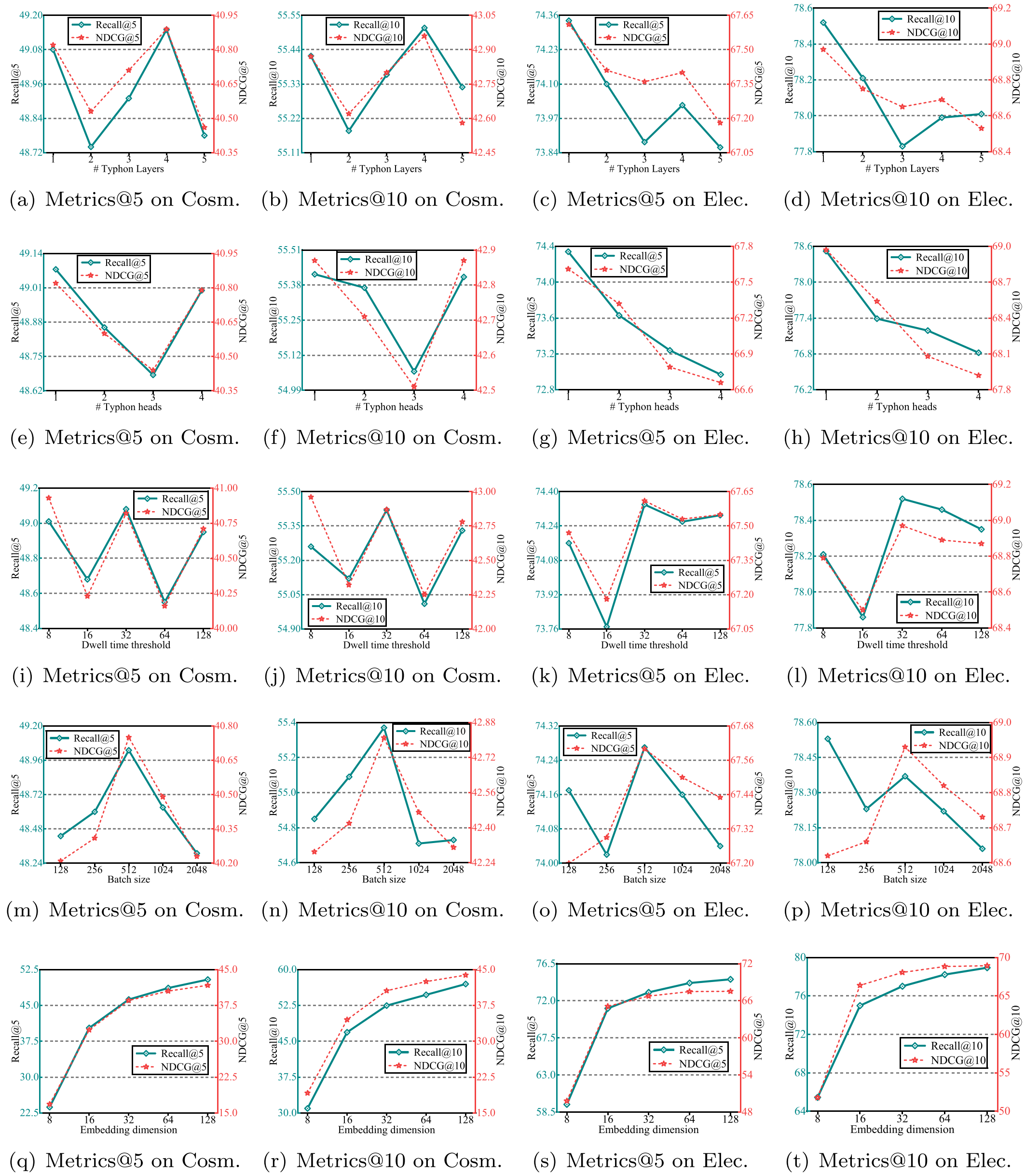
For the number of Typhon layers, the *Cosmetics* dataset achieves the best performance with four layers, suggesting that complex cosmetic preferences benefit from deeper feature transformations. In contrast, the *Electronics Store* dataset performs best with a single layer, indicating that a simpler architecture is sufficient. This difference likely reflects distinct decision processes: cosmetic purchases involve more nuanced aesthetic judgments, whereas electronics purchases are driven primarily by functional attributes.

Regarding attention heads, single-head Typhon achieves the best results on both datasets, with performance degrading as the number of heads increases. This observation challenges the common assumption that multiple heads are necessary to capture diverse relational aspects [118], and instead suggests that recommendation tasks may benefit from more focused attention.

The dwell time threshold analysis reveals a consistent optimum at 32 across both datasets, with performance declining for both smaller and larger thresholds. This value appears to approximate the typical engagement window during which users meaningfully process product information before interacting. Shorter thresholds may truncate genuine browsing behavior, whereas longer thresholds risk incorporating intervals of inactivity or distraction. This observation that performance varies with dwell time threshold demonstrates that cognitive foraging costs depend on decision context, a nuance typically overlooked by heuristic feature engineering approaches. For example, across different product categories (e.g., high-involvement electronics versus low-involvement groceries), the optimal dwell time threshold is expected to differ, as users allocate varying levels of cognitive effort when evaluating different types of products. Additionally, our model maintains superior performance relative to most of baselines across the entire tested range of thresholds, demonstrating the modeling of explicit foraging costs is robust to reasonable variations.

Batch size influences the model's optimization stability and generalization performance. Results show a consistent inverted U-shaped performance trend across both datasets, with optimal performance achieved at the batch size 512. Overly small batch sizes lead to high variance in gradient estimation and unstable training dynamics, while excessively large batch sizes reduce the model's generalization ability and dilute fine-grained sequential user foraging signals, both resulting in moderate performance degradation.

As a core operationalization decision, embedding dimension governs the model's ability to capture multi-faceted foraging-related signals. The results reveal consistent performance trends across both datasets, with all evaluation metrics improving as the embedding dimension increases. This pattern aligns with the fundamental role of latent embeddings, as an overly small embedding size severely restricts the model's representational capacity, preventing it from effectively encoding diverse information foraging signals. Additionally, for the



(a) Metrics@5 on Cosm. (b) Metrics@10 on Cosm. (c) Metrics@5 on Elec. (d) Metrics@10 on Elec.

(e) Metrics@5 on Cosm. (f) Metrics@10 on Cosm. (g) Metrics@5 on Elec. (h) Metrics@10 on Elec.

(i) Metrics@5 on Cosm. (j) Metrics@10 on Cosm. (k) Metrics@5 on Elec. (l) Metrics@10 on Elec.

(m) Metrics@5 on Cosm. (n) Metrics@10 on Cosm. (o) Metrics@5 on Elec. (p) Metrics@10 on Elec.

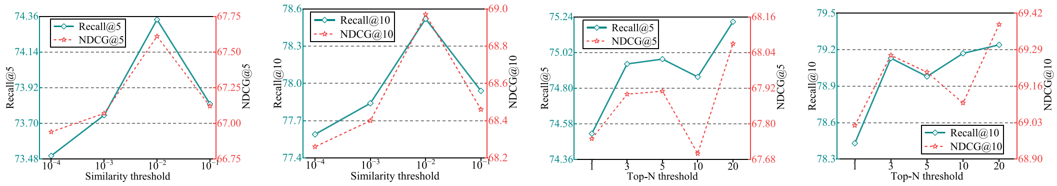
(q) Metrics@5 on Cosm. (r) Metrics@10 on Cosm. (s) Metrics@5 on Elec. (t) Metrics@10 on Elec.

Fig. 7. Tiresias performance with varying parameter settings on the two datasets.

higher-sparsity Electronics Store dataset, an earlier onset of diminishing marginal returns is observed. One key reason is that the limited pool of interaction signals fails to support further gains from increased representational capacity, while oversized embeddings introduce additional computational overhead and a higher risk of noise encoding.

For the user affinity network, sensitivity analyses are conducted on both the similarity threshold and graph construction method. In the main setup, affinitive relationships are derived from user–product interactions using Jaccard similarity, with the top 1% of relationships retained to filter out low-confidence edges. The similarity threshold is varied across {0.1, 0.01, 0.001, 0.0001}, corresponding to retaining the top 10%, 1%, 0.1%, and 0.01% of user relationships, respectively. In addition, an alternative per-user Top-N neighborhood construction method is adopted, where the top-N most similar users are retained for each user based on the same Jaccard similarity. The value of N is tuned over {1, 3, 5, 10, 20}. The results are reported in Fig. 8.

In the first group of experiments, the model achieves its best performance at a similarity threshold of 0.01, with performance declining at both higher and lower thresholds. This pattern reflects a trade-off between signal quality and graph connectivity. Lenient thresholds (e.g., 0.1) introduce many weak and non-informative edges, diluting the information scent provided by genuinely similar users and reducing the model's ability to capture preferences accurately. In contrast, overly strict thresholds (e.g., 0.0001) excessively prune the graph, removing meaningful affinitive relationships and limiting access to complementary peer-derived cues. In the second group of experiments, recommendation performance improves consistently as N increases. Larger neighborhoods expose users to more affinitive relationships and richer peer-derived information scent, helping alleviate interaction sparsity and enhancing the model's understanding of evolving preferences. However, increasing N also densifies the user affinity network and raises computational costs. Notably, the model outperforms most baselines even under suboptimal threshold settings, suggesting that peer-derived proximal cues remain valuable without extensive parameter tuning. The results further indicate that sufficiently similar peers provide reliable signals about valuable foraging opportunities.



(a) Metrics@5 on Cosm. (b) Metrics@10 on Cosm. (c) Metrics@5 on Elec. (d) Metrics@10 on Elec.

Fig. 8. Tiresias performance with varying user affinity network configurations on Electronic Store.

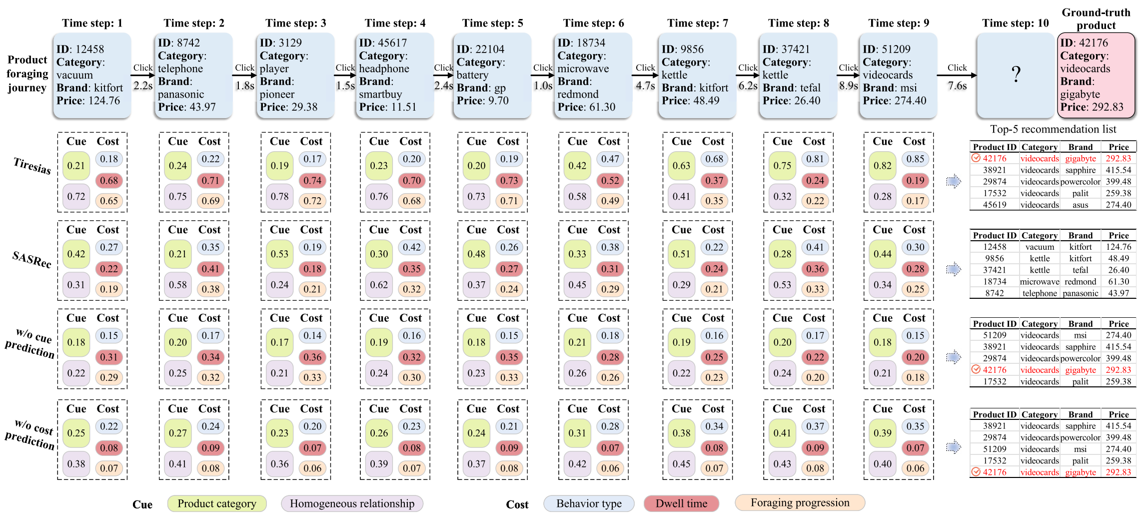


Fig. 9. A product foraging journey instance from the Electronic Store dataset.

4.7. Post-hoc analysis on exemplar instances

This section presents a post-hoc analysis to illustrate representation-level patterns consistent with the evolution of users' dynamic preferences across product foraging journeys within our IFT-guided framework. Specifically, SASRec, Tiresias, and two ablated variants of Tiresias are implemented on the Electronic Store dataset. A representative case is presented in which Tiresias successfully ranks the interacted product within the top-5, while the competing models fail. Fig. 9 shows, at each time step, the cosine similarity between the learned user preference representation and both proximal cue and foraging cost features is computed. The results reveal a two-phase foraging pattern: an initial exploration phase with cross-category browsing (time steps 1–5), followed by an exploitation phase with preference convergence to computer components and kitchen appliances (time steps 6–9).

The top-5 recommendation lists further demonstrate the effectiveness of the IFT-guided design. Tiresias ranks the user's actual next product first, whereas SASRec fails to include it in the top-5 list. The two ablated variants rank it only fourth and fifth, respectively. For Tiresias, the learned preference representation aligns most strongly with the affinitive relationship cue during exploration, suggesting that users rely on peer-derived signals under high preference uncertainty. Alignment with foraging cost features also remains high, reflecting elevated cognitive costs during exploratory search. As the user transitions to exploitation, similarity between the preference representation and the product category cue increases sharply from 0.20 to 0.82, indicating a shift from external to internal cues as preferences become more defined.

As a baseline model, SASRec fails to reveal any meaningful patterns, with cosine similarities between its learned preference representation and all cue/cost features fluctuating randomly across the user foraging journey. This suggests that cognition-agnostic models primarily fit observed interaction patterns, and their learned representations fail to align with the proximal cues and foraging costs. Ablation experiments further support the necessity of the cue and cost prediction tasks. Removing the cue prediction task eliminates any meaningful alignment between the learned preference representation and the two proximal cue features across the full journey. Similarly, excluding the cost prediction task impairs the model's ability to learn representative foraging cost features, with dwell time and foraging progression cost features consistently showing low similarity to user preferences. These results show that the cue and cost prediction tasks explicitly guide the model to learn representations that align with IFT-based cue and cost semantics, facilitating a comprehensive understanding of behavior patterns in users' product foraging journey.

In Section 4.5, users are segmented into three groups based on foraging trail length. The cold-start group constitutes the largest proportion of users, posing a challenge for reducing premature churn. To further investigate user exit and sustained engagement, a post-hoc analysis is conducted using the same grouping. Cold-start users exhibit the shortest observed foraging journeys, which may be interpreted as a greater tendency toward early disengagement, whereas active users show the longest observed foraging journeys, which may be interpreted as a greater tendency toward sustained engagement. For each group, the mean cosine similarity between user preference representations and fused cue and cost features is computed over the first five time steps. The fused cue feature is defined as the element-wise sum of affinitive relationship and product category features, while the fused cost feature combines behavior type, dwell time, and foraging progression features. The resulting alignment patterns across user groups are presented in Fig. 10.

For the cold-start group, cost-preference similarity consistently exceeds cue-preference similarity across all five time steps, and the gap widens over time. This pattern supports the IFT perspective that foraging costs exert a stronger influence than proximal cues on preference formation, potentially leading to shorter foraging trails. In contrast,

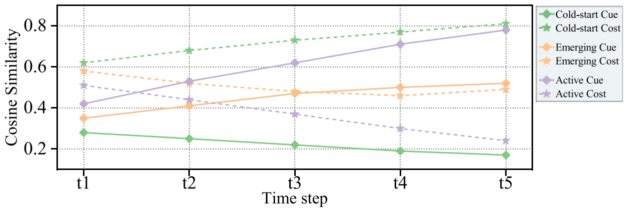


Fig. 10. Mean cue-preference and cost-preference similarity across three user groups.

for the active group, cue-preference similarity surpasses cost-preference similarity from the second time step onward, with the difference increasing throughout the session. This suggests that proximal cues play a dominant role in shaping preferences among users who maintain sustained engagement. For the emerging group, cue-preference and cost-preference similarities gradually converge, indicating a more balanced decision-making process that does not consistently develop into sustained foraging behavior.

4.8. IFT-derived predictions and evidence

In the product foraging context, IFT suggests two testable predictions to explain user foraging behavior. First, users evaluate product patches by jointly considering proximal cues and foraging costs, implying that both signal types contain decision-relevant information. Second, product foraging is inherently dynamic. During the early stages of a foraging journey, when uncertainty is high, users are expected to be more sensitive to foraging costs. If perceived utility falls below incurred costs, users may terminate their search prematurely; otherwise, they may increasingly rely on proximal cues to evaluate product attractiveness as the session progresses. Our findings support both predictions, although to varying degrees. The ablations directly show that cue/cost features and auxiliary tasks improve model performance, providing theory-consistent support for the IFT interpretation. Support for the second prediction is more indirect. Results from the foraging trail length experiments and post hoc analyses reveal distinct patterns of cue and cost sensitivity across user groups and journey stages, consistent with the dynamic adaptation mechanisms proposed by IFT.

5. Discussion

Although online e-commerce platforms provide access to vast amounts of product information, they also increase the cognitive effort required for purchase decisions. To address this challenge, we propose Tiresias, a theory-driven recommender system grounded in IFT. Tiresias incorporates an IFT-guided feature selection strategy that distinguishes proximal cues from foraging costs, a preference extractor that models the evolving nature of product foraging behavior, and a multi-task learning objective that aligns optimization with users' cognitive decision processes. Extensive experiments on two real-world datasets and 30 representative baselines demonstrate the effectiveness of Tiresias. Ablation studies and supplementary analyses further validate the contributions of the theory-guided design, multi-task objective, and model architecture. Beyond improving recommendation accuracy, our findings show that integrating IFT into recommender system design provides a principled framework for capturing the cognitive mechanisms underlying user decision-making.

5.1. Research contributions

Our study makes three contributions. First, it contributes to the decision support systems literature by translating IFT from an ex post explanatory theory of user behavior into an actionable framework for cognitive-aware decision support systems. Existing recommender systems, including state-of-the-art models, typically treat decision-making as a black box, emphasizing correlations between user behavior and next-product interactions rather than the cognitive mechanisms underlying them. We ground the design of our recommender system's three core modules in IFT, aligning behavior prediction with users' cognitive processes and providing a generalizable paradigm for cognitive-aware decision support systems. Notably, the study clarifies that the generalizable design knowledge contributed by this work is the coherent IFT-guided design paradigm, rather than discrete technical components. Specific implementation choices can be freely replaced while preserving the core theoretical framework. Second, this study contributes to the recommender systems literature by bridging multiple recommendation domains through a unified feature-selection perspective. Our framework integrates insights from sequential, temporal, multi-behavior, attribute-aware, and social recommendation. This approach moves beyond the siloed development of these domains and offers a coherent framework that captures the multifaceted nature of user decision-making. Third, this study contributes to the design science literature by introducing IFT-grounded design innovations across the three core modules of recommender system development. Specifically, we propose an IFT-guided feature-selection approach that categorizes features into proximal cues and foraging costs, providing a principled alternative to heuristic methods. We also develop an IFT-guided user preference extractor to model product foraging behavior and a multi-task learning objective that aligns optimization with users' cognitive processes of cue evaluation and cost assessment during foraging decisions.

5.2. Practical implications

For platform and system designers, our findings suggest three IFT-aligned design principles. First, designers should adopt IFT-guided feature selection instead of heuristic approaches that rely solely on product IDs or indiscriminately include all available features. This strategy aligns the input space with the factors driving user decisions, improving recommendation performance. Second, preference learning should prioritize sequential modeling over static user representations. Although graph-based methods are widely used, they often fail to capture evolving preferences. Sequential modeling better tracks preference dynamics across interactions, enabling context-aware recommendations that reflect users' current decision states. Third, designers should employ multi-task learning objectives that align optimization with users' cognitive processes rather than relying solely on next-product prediction. In our framework, auxiliary cue- and cost-prediction tasks encourage the model to learn the cost–benefit trade-offs underlying product-foraging decisions.

Additional design principles emerge from the post-hoc analyses. The results reveal a cost-dominant decision pattern among cold-start users, with the gap between cost-preference similarity and cue-preference similarity widening over time. For these users, the cognitive costs of continued foraging outweigh the informational value of proximal cues. Accordingly, practitioners should simplify interaction paths, reduce unnecessary decision steps, and minimize cognitive friction, as improvements in cue relevance alone are unlikely to overcome high foraging costs. Moreover, across all user groups, the early stages of a foraging journey are characterized by greater sensitivity to foraging costs. Designers should therefore prioritize cost reduction during initial interactions and gradually shift attention to cue refinement as the session progresses.

Our findings also reveal boundary conditions for these design principles. First, the IFT-guided framework is most valuable in high-uncertainty contexts characterized by long decision cycles and information overload. For low-involvement products, where decisions are more habitual and involve limited information search, designers may reduce the emphasis on foraging cost signals. Second, the advantage of

Tiresias is smaller for emerging users, who exhibit a balanced cue–cost decision pattern. For this group, practitioners should jointly optimize cue relevance and foraging costs. Finally, parameter sensitivity analyses indicate strong robustness but identify two configurations that reduce performance: excessively small embedding dimensions, which fail to capture fine-grained cue and cost information, and too many attention heads in the Typhon module, which fragment attention and weaken alignment with information scent assessment and cost evaluation.

5.3. Limitations and future work

Our evaluation is limited to offline ranking performance and post-hoc representation analysis, which imposes boundaries on the conclusions that can be drawn. To avoid overinterpretation, we align the conclusions with the available evidence as follows. First, the main directly showed conclusions include the superior offline ranking performance of our model and the contribution of each core component demonstrated via ablation studies. Additionally, the indirectly suggested conclusion is the learned representations are consistent with cue and cost dynamics. Finally, the conclusions that are not directly shown include actual reductions in cognitive burden for users, improved decision quality in real-world shopping scenarios, and changes in sustained engagement or abandonment rates in real systems. To address these limitations, future research can focus on three directions aligned with our evidence boundaries. First, future work could explore a broader spectrum of features, for example incorporating product images as additional proximal cues and price as an additional foraging cost. Second, employing experimental methods, such as eye-tracking and cognitive load measurements, to validate the hypothesized cognitive mechanisms. Third, conducting production A/B testing complemented by user surveys to measure real-world behavioral and cognitive outcomes.

CRediT authorship contribution statement

Weiyue Li: Writing – original draft, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ming Gao:** Writing – review & editing, Supervision, Resources. **Jingmin An:** Writing – review & editing. **Bowei Chen:** Writing – review & editing. **Jiafu Tang:** Writing – review & editing. **Xuhui Wang:** Writing – review & editing. **Yeming Gong:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported and funded by the National Natural Science Foundation of China (No. 72293563, 72442025, 72501053); the Basic Scientific Research Project of Liaoning Provincial Department of Education (No. JYTZD2023050); the Natural Science Foundation of Liaoning Province (No. 2024-MS-175); and the Dalian Scientific and Technological Talents Innovation Support Plan (No. 2022RG17). The authors would like to extend sincere gratitude to Zilong Liu and Maoxin (Molson) Han from the School of Management Science and Engineering at Dongbei University of Finance and Economics for their valuable discussions and insightful comments on this work. We also wish to thank Hongyu Wang of the Institute of Computing Technology, Chinese Academy of Sciences for his technical support and expertise.

Data availability

Data will be made available on request.

References

[1] M. Schrage, The transformational power of recommendation, MIT Sloan Manag. Rev. 62 (2) (2021) 17–21.
[2] Z. Li, C. Yang, Y. Chen, X. Wang, H. Chen, G. Xu, L. Yao, M. Sheng, Graph and sequential neural networks in session-based recommendation: A survey, ACM Comput. Surv. 57 (2) (2025) 40:1–40:37.
[3] K. Kim, S. Kim, G. Lee, J. Jung, K. Shin, Multi-behavior recommender systems: A survey, in: PAKDD, in: Lecture Notes in Computer Science, vol. 15873, 2025, pp. 435–452.
[4] X. Zhang, B. Xu, C. Li, B. He, H. Lin, C. Ma, F. Ma, A survey on side information-driven session-based recommendation: From a data-centric perspective, IEEE Trans. Knowl. Data Eng. 37 (8) (2025) 4411–4431.
[5] V. Bogina, T. Kuflik, D. Jannach, M. Bieliková, M. Kompan, C. Trattner, Considering temporal aspects in recommender systems: A survey, User Model. User Adapt. Interact. 33 (1) (2023) 81–119.
[6] K. Sharma, Y. Lee, S. Nambi, A. Salian, S. Shah, S. Kim, S. Kumar, A survey of graph neural networks for social recommender systems, ACM Comput. Surv. 56 (10) (2024) 265.
[7] D.L. R., N. Pervin, Towards generating scalable personalized recommendations: Integrating social trust, social bias, and geo-spatial clustering, Decis. Support Syst. 122 (2019).
[8] J. Liao, W. Zhou, F. Luo, J. Wen, M. Gao, X. Li, J. Zeng, SocialLGN: Light graph convolution network for social recommendation, Inf. Sci. 589 (2022) 595–607.
[9] W. Li, M. Gao, B. Chen, J. An, Y. Gong, Social capital matters: Towards comprehensive user preference for product recommendation with deep learning, Decis. Support Syst. 198 (2025) 114527.
[10] A. Ramaprasad, Cognitive process as a basis for MIS and DSS design, Manag. Sci. 33 (2) (1987) 139–148.
[11] L. Wang, X. Zhao, N. Liu, Z. Shen, C. Zou, Cognitive process-driven model design: A deep learning recommendation model with textual review and context, Decis. Support Syst. 176 (2024) 114062.
[12] W. Kang, J.J. McAuley, Self-attentive sequential recommendation, in: ICDM, 2018, pp. 197–206.
[13] X. Wang, H. Ji, C. Shi, B. Wang, Y. Ye, P. Cui, P.S. Yu, Heterogeneous graph attention network, in: WWW, 2019, pp. 2022–2032.
[14] C. Liu, X. Li, G. Cai, Z. Dong, H. Zhu, L. Shang, Noninvasive self-attention for side information fusion in sequential recommendation, in: AAAI, 2021, pp. 4249–4256.
[15] K. Peffers, T. Tuunanen, M.A. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, J. Manage. Inf. Syst. 24 (3) (2008) 45–77.
[16] K.R. Larsen, R. Lukyanenko, R.M. Mueller, V.C. Storey, J. Parsons, D.E. VanderMeer, D.S. Hovorka, Validity in design science, MIS Q. 49 (4) (2025) 1267–1294.
[17] J. Lawrance, C. Bogart, M.M. Burnett, R.K.E. Bellamy, K. Rector, S.D. Fleming, How programmers debug, revisited: An information foraging theory perspective, IEEE Trans. Softw. Eng. 39 (2) (2013) 197–215.
[18] S.D. Fleming, C. Scaffidi, D. Piorkowski, M.M. Burnett, R.K.E. Bellamy, J. Lawrance, I. Kwan, An information foraging theory perspective on tools for debugging, refactoring, and reuse tasks, ACM Trans. Softw. Eng. Methodol. 22 (2) (2013) 14:1–14:41.
[19] L. Azzopardi, P. Thomas, N. Craswell, Measuring the utility of search engine result pages: An information foraging based measure, in: SIGIR, 2018, pp. 605–614.
[20] T. Schnabel, P.N. Bennett, T. Joachims, Shaping feedback data in recommender systems with interventions based on information foraging theory, in: WSDM, 2019, pp. 546–554.
[21] T. Liang, Y. Yang, D. Chen, Y. Ku, A semantic-expansion approach to personalized knowledge recommendation, Decis. Support Syst. 45 (3) (2008) 401–412.
[22] Y. Wang, W. Dai, Y. Yuan, Website browsing aid: A navigation graph-based recommendation system, Decis. Support Syst. 45 (3) (2008) 387–400.
[23] S. Kim, W. Shin, H. Kim, Predicting online customer purchase: The integration of customer characteristics and browsing patterns, Decis. Support Syst. 177 (2024) 114105.
[24] X. Yi, L. Hong, E. Zhong, N.N. Liu, S. Rajan, Beyond clicks: Dwell time for personalization, in: RecSys, 2014, pp. 113–120.
[25] H. Zhao, L. Zhang, J. Xu, G. Cai, Z. Dong, J. Wen, Uncovering user interest from biased and noised watch time in video recommendation, in: RecSys, 2023, pp. 528–539.
[26] G. Tang, X. Zhu, J. Guo, S. Dietze, Time enhanced graph neural networks for session-based recommendation, Knowl.-Based Syst. 251 (2022) 109204.
[27] V. Bogina, T. Kuflik, Incorporating dwell time in session-based recommendations with recurrent neural networks, in: RecSys, 2017, pp. 57–59.
[28] O. Mokryn, V. Bogina, T. Kuflik, Will this session end with a purchase? Inferring current purchase intent of anonymous visitors, Electron. Commer. Res. Appl. 34 (2019).

[29] Y. Liu, Y. Chen, P. Chang, A deep multi-embedding model for mobile application recommendation, Decis. Support Syst. 173 (2023) 114011.
[30] Y. Lai, S. Huang, C. Chen, C. Wu, Multi-interest transfer using contrastive learning for cross-domain recommendation, Decis. Support Syst. 195 (2025) 114473.
[31] W. Zhang, R. Xie, P. Quan, Z. Ma, Product return prediction in live streaming e-commerce with cross-modal contrastive transformer, Decis. Support Syst. 194 (2025) 114470.
[32] L. Yu, W. Gong, D. Zhang, Live streaming channel recommendation based on viewers' interaction behavior: A hypergraph approach, Decis. Support Syst. 184 (2024) 114272.
[33] P. Pirolli, S. Card, Information foraging, Psychol Rev 106 (4) (1999) 643.
[34] D. Piorkowski, S.D. Fleming, C. Scaffidi, C. Bogart, M.M. Burnett, B.E. John, R.K.E. Bellamy, C. Swart, Reactive information foraging: An empirical investigation of theory-based recommender systems for programmers, in: CHI, 2012, pp. 1471–1480.
[35] J. Liu, S. Zhang, J. Yang, Characterizing web usage regularities with information foraging agents, IEEE Trans. Knowl. Data Eng. 16 (5) (2004) 566–584.
[36] C. Yi, Z.J. Jiang, I. Benbasat, Designing for diagnosticity and serendipity: An investigation of social product-search mechanisms, Inf. Syst. Res. 28 (2) (2017) 413–429.
[37] B. Wang, J. Liu, Understanding users' dynamic perceptions of search gain and cost in sessions: An expectation confirmation model, J. Assoc. Inf. Sci. Technol. 75 (9) (2024) 937–956.
[38] F. Loumakis, S. Stumpf, D. Grayson, This image smells good: Effects of image information scent in search engine results pages, in: CIKM, 2011, pp. 475–484.
[39] N. Pervin, A. Kulkarni, A. Adarsh, S. Som, Knowledge-based context-aware group recommender system for point of interest recommendation, Decis. Support Syst. 196 (2025) 114485.
[40] D.M. Luu, E. Lim, Do your friends make you buy this brand? - Modeling social recommendation with topics and brands, Data Min. Knowl. Discov. 32 (2) (2018) 287–319.
[41] A.R. Hevner, S.T. March, J. Park, S. Ram, Design science in information systems research, MIS Q. 28 (1) (2004) 75–105.
[42] P. Kumar, D. Javeed, A.K.M.N. Islam, X.R. Luo, DeepSecure: A computational design science approach for interpretable threat hunting in cybersecurity decision making, Decis. Support Syst. 188 (2025) 114351.
[43] X. Wang, Y. Feng, S. Wang, D. Wang, T.C.E. Cheng, An explainable lesion detection transformer model for medical imaging diagnosis decision support: Design science research, Decis. Support Syst. 196 (2025) 114492.
[44] C. Anderson, P. Shrestha, S. Bhunia, A. Carvalho, Y. Lee, Blockchain-based token system for incentivizing peer review: A design science approach, Decis. Support Syst. 197 (2025) 114514.
[45] Z. Liu, J. Li, X. Wang, Y. Guo, How search and evaluation cues influence consumers' continuous watching and purchase intentions: An investigation of live-stream shopping from an information foraging perspective, J. Bus. Res. 168 (2023) 114233.
[46] W.W. Moe, Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream, J. Consum. Psychol. 13 (2003) 29–39.
[47] A.G. Close, M. Kukar-Kinney, Beyond buying: Motivations behind consumers' online shopping cart use, J. Bus. Res. 63 (2010) 986–992.
[48] C. Groening, J. Wiggins, I. Raoofpanah, Wish list thinking: The quasi-endowment effect's impact on online wish lists outcomes, J. Consum. Behav. 20 (2) (2021) 412–425.
[49] G. Fisher, Measuring the factors influencing purchasing decisions: Evidence from cursor tracking and cognitive modeling, Manag. Sci. 69 (8) (2023) 4558–4578.
[50] S. Bashir, S. Anwar, Z. Awan, T.W. Qureshi, A.B. Memon, A holistic understanding of the prospects of financial loss to enhance shopper's trust to search, recommend, speak positive and frequently visit an online shop, J. Retail. Consum. Serv. 42 (2018) 169–174.
[51] F. Reshadi, M.P. Fitzgerald, The pain of payment: A review and research agenda, Psychol. Mark. 40 (2023) 1672–1688.
[52] M. Guan, J. Li, Y. Zhang, Y. Liu, Coparticipant effect in group buying: How coparticipant response speeds affect consumer postpurchase regret? Decis. Support Syst. 171 (2023) 113980.
[53] H. Zhao, X. Wang, L. Jiang, To purchase or to remove? Online shopping cart warning pop-up messages can polarize liking and purchase intention, J. Bus. Res. 132 (2021) 813–836.
[54] M. O'Neill, A. Palmer, Cognitive dissonance and the stability of service quality perceptions, J. Serv. Mark. 18 (6) (2004) 433–449.
[55] P. Yin, P. Luo, W. Lee, M. Wang, Silence is also evidence: interpreting dwell time for recommendation from psychological perspective, in: SIGKDD, 2013, pp. 989–997.
[56] G.L. Lohse, S. Bellman, E.J. Johnson, Consumer buying behavior on the internet: Findings from panel data, J. Interact. Mark. 14 (1) (2000) 15–29.
[57] W. Hong, J.Y.L. Thong, K.Y. Tam, The effects of information format and shopping task on consumers' online shopping behavior: A cognitive fit perspective, J. Manage. Inf. Syst. 21 (3) (2004) 149–184.
[58] J. Mosteller, N. Donthu, S. Eroglu, The fluent online shopping experience, J. Bus. Res. 67 (11) (2014) 2486–2493.
[59] P.A.L. Mazursky;, A longitudinal assessment of consumer satisfaction/dissatisfaction: The dynamic aspect of the cognitive process, J. Mark. Res. 20 (4) (1983) 393–404.
[60] X. Li, C.K. Hsee, The psychology of marginal utility, J. Consum. Res. 48 (1) (2021) 169–188.
[61] J. Murphy, C.F. Hofacker, R. Mizerski, Primacy and recency effects on clicking behavior, J. Comput. Mediat. Commun. 11 (2) (2006) 522–535.
[62] J.A. Rosa, J.F. Porac, Categorization bases and their influence on product category knowledge structures, Psychol. Mark. 19 (6) (2002) 503–531.
[63] Á. Garrido-Morgado, Ó. González-Benito, M. Martos-Partal, K. Campo, Which products are more responsive to in-store displays: Utilitarian or hedonic? J. Retail. 97 (3) (2021) 477–491.
[64] X. Shen, K.Z.K. Zhang, S.J. Zhao, Herd behavior in consumers' adoption of online reviews, J. Assoc. Inf. Sci. Technol. 67 (11) (2016) 2754–2765.
[65] A.W. Ding, S. Li, Herding in the consumption and purchase of digital goods and moderators of the herding bias, J. Acad. Mark. Sci. 47 (3) (2019) 460–478.
[66] Y. Chen, Herd behavior in purchasing books online, Comput. Hum. Behav. 24 (5) (2008) 1977–1992.
[67] J.-H. Huang, Y.-F. Chen, Herding in online product choice, Psychol. Mark. 23 (5) (2006) 413–428.
[68] M. Ruef, H.E. Aldrich, N.M. Carter, The structure of founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs, Am. Sociol. Rev. 68 (2) (2003) 195–222.
[69] E. Ismagilova, E. Slade, N.P. Rana, Y.K. Dwivedi, The effect of characteristics of source credibility on consumer behaviour: A meta-analysis, J. Retail. Consum. Serv. 53 (2020) 101736.
[70] M. Gupta, S. Gupta, G.K. Tayi, GASP: A graph augmentation-based approach for sign prediction of ties in social networks, J. Assoc. Inf. Syst. 26 (5) (2025) 1.
[71] A. Grover, J. Leskovec, Node2vec: Scalable feature learning for networks, in: SIGKDD, 2016, pp. 855–864.
[72] J. Su, M.H.M. Ahmed, Y. Lu, S. Pan, W. Bo, Y. Liu, RoFormer: Enhanced transformer with rotary position embedding, Neurocomputing 568 (2024) 127063.
[73] X. Zhang, X. Chang, M. Li, A.K. Roy-Chowdhury, J. Chen, S. Oymak, Selective attention: Enhancing transformer through principled context control, in: NeurIPS, 2024, pp. 11061–11086.
[74] T. Ye, L. Dong, Y. Xia, Y. Sun, Y. Zhu, G. Huang, F. Wei, Differential transformer, in: ICLR, 2025, pp. 1–21.
[75] X. Yang, X. Wang, Kolmogorov-arnold transformer, in: ICLR, 2025, pp. 1–24.
[76] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljacic, T.Y. Hou, M. Tegmark, KAN: Kolmogorov-arnold networks, in: ICLR, 2025, pp. 1–48.
[77] A. Molina, P. Schramowski, K. Kersting, Padé activation units: End-to-end learning of flexible activation functions in deep networks, in: ICLR, 2020, pp. 1–17.
[78] H. Wang, S. Ma, L. Dong, S. Huang, D. Zhang, F. Wei, DeepNet: Scaling transformers to 1,000 layers, IEEE Trans. Pattern Anal. Mach. Intell. 46 (10) (2024) 6761–6774.
[79] J. Zhu, X. Chen, K. He, Y. LeCun, Z. Liu, Transformers without normalization, in: CVPR, 2025, pp. 14901–14911.
[80] W. Krichene, S. Rendle, On sampled metrics for item recommendation, Commun. ACM 65 (7) (2022) 75–83.
[81] W.X. Zhao, S. Mu, Y. Hou, Z. Lin, Y. Chen, X. Pan, K. Li, Y. Lu, H. Wang, C. Tian, Y. Min, Z. Feng, X. Fan, X. Chen, P. Wang, W. Ji, Y. Li, X. Wang, J. Wen, RecBole: Towards a unified, comprehensive and efficient framework for recommendation algorithms, in: CIKM, 2021, pp. 4653–4664.
[82] W.X. Zhao, Y. Hou, X. Pan, C. Yang, Z. Zhang, Z. Lin, J. Zhang, S. Bian, J. Tang, W. Sun, Y. Chen, L. Xu, G. Zhang, Z. Tian, C. Tian, S. Mu, X. Fan, X. Chen, J. Wen, RecBole 2.0: Towards a more up-to-date recommendation library, in: CIKM, ACM, 2022, pp. 4722–4726.
[83] I. Loshchilov, F. Hutter, Decoupled weight decay regularization, in: ICLR, 2019, pp. 1–19.
[84] X. Ren, L. Xia, Y. Yang, W. Wei, T. Wang, X. Cai, C. Huang, SSLRec: A self-supervised learning framework for recommendation, in: WSDM, 2024, pp. 567–575.
[85] F. Aiolli, Efficient top-n recommendation for very large scale binary rated datasets, in: RecSys, 2013, pp. 273–280.
[86] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, M. Wang, LightGCN: Simplifying and powering graph convolution network for recommendation, in: SIGIR, 2020, pp. 639–648.
[87] H. Wei, J. Wang, Y. Ji, M. Guang, C. Yan, Hierarchical neighbor-enhanced graph contrastive learning for recommendation, Knowl.-Based Syst. 315 (2025) 113263.
[88] B. Hidasi, A. Karatzoglou, L. Baltrunas, D. Tikk, Session-based recommendations with recurrent neural networks, in: ICLR, 2016, pp. 1–10.
[89] J. Li, P. Ren, Z. Chen, Z. Ren, T. Lian, J. Ma, Neural attentive session-based recommendation, in: CIKM, ACM, 2017, pp. 1419–1428.
[90] I. Baek, M. Yoon, S. Park, J. Lee, MUFFIN: Mixture of user-adaptive frequency filtering for sequential recommendation, 2025, CoRR, abs/2508.13670, arXiv: 2508.13670.

[91] Z. Li, A. Sun, C. Li, DiffuRec: A diffusion model for sequential recommendation, ACM Trans. Inf. Syst. 42 (3) (2024) 66:1–66:28.
[92] T. Chen, R.C. Wong, An efficient and effective framework for session-based social recommendation, in: WSDM, 2021, pp. 400–408.
[93] W. Song, Z. Xiao, Y. Wang, L. Charlin, M. Zhang, J. Tang, Session-based social recommendation via dynamic graph attention networks, in: WSDM, 2019, pp. 555–563.
[94] J. Gao, X. Zhao, M. Li, M. Zhao, R. Wu, R. Guo, Y. Liu, D. Yin, SMLP4Rec: An efficient all-MLP architecture for sequential recommendations, ACM Trans. Inf. Syst. 42 (3) (2024) 86:1–86:23.
[95] F. Sun, J. Liu, J. Wu, C. Pei, X. Lin, W. Ou, P. Jiang, BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer, in: CIKM, 2019, pp. 1441–1450.
[96] X. Xie, F. Sun, Z. Liu, S. Wu, J. Gao, J. Zhang, B. Ding, B. Cui, Contrastive learning for sequential recommendation, in: ICDE, 2022, pp. 1259–1273.
[97] P. Zhou, Q. Ye, Y. Xie, J. Gao, S. Wang, J.B. Kim, C. You, S. Kim, Attention calibration for transformer-based sequential recommendation, in: CIKM, 2023, pp. 3595–3605.
[98] Z. Pan, F. Cai, W. Chen, H. Chen, M. de Rijke, Star graph neural networks for session-based recommendation, in: CIKM, 2020, pp. 1195–1204.
[99] Y. Cao, L. Yang, Z. Liu, Y. Liu, C. Wang, Y. Liang, H. Peng, P.S. Yu, Graph-sequential alignment and uniformity: Toward enhanced recommendation systems, in: WWW, 2025, pp. 888–892.
[100] X. Lin, J. Luo, J. Pan, W. Pan, Z. Ming, X. Liu, S. Huang, J. Jiang, Multi-sequence attentive user representation learning for side-information integrated sequential recommendation, in: WSDM, 2024, pp. 414–423.
[101] P. Zhou, J. Gao, Y. Xie, Q. Ye, Y. Hua, J. Kim, S. Wang, S. Kim, Equivariant contrastive learning for sequential recommendation, in: RecSys, 2023, pp. 129–140.
[102] J. Li, Y. Wang, J.J. McAuley, Time interval aware self-attention for sequential recommendation, in: WSDM, 2020, pp. 322–330.
[103] W. Xiao, H. Wang, Q. Zhou, Q. Wang, SS4Rec: Continuous-time sequential recommendation with state space models, 2025, CoRR, abs/2502.08132, arXiv: 2502.08132.
[104] H. Xuan, Y. Liu, B. Li, H. Yin, Knowledge enhancement for contrastive multi-behavior recommendation, in: WSDM, 2023, pp. 195–203.
[105] A. García-Durán, A. Bordes, N. Usunier, Effective blending of two and three-way interactions for modeling multi-relational data, in: ECML-PKDD, in: Lecture Notes in Computer Science, vol. 8724, 2014, pp. 434–449.
[106] Y. Lin, Z. Liu, M. Sun, Y. Liu, X. Zhu, Learning entity and relation embeddings for knowledge graph completion, in: AAAI, 2015, pp. 2181–2187.
[107] W. Wei, C. Huang, L. Xia, Y. Xu, J. Zhao, D. Yin, Contrastive meta learning with behavior multiplicity for recommendation, in: WSDM, 2022, pp. 1120–1128.
[108] X. Xin, X. Liu, H. Wang, P. Ren, Z. Chen, J. Lei, X. Shi, H. Luo, J.M. Jose, M. de Rijke, Z. Ren, Improving implicit feedback-based recommendation through multi-behavior alignment, in: SIGIR, 2023, pp. 932–941.
[109] Y. Yang, C. Huang, L. Xia, Y. Liang, Y. Yu, C. Li, Multi-behavior hypergraph-enhanced transformer for sequential recommendation, in: SIGKDD, 2022, pp. 2263–2274.
[110] Z. Liu, S. Mei, C. Xiong, X. Li, S. Yu, Z. Liu, Y. Gu, G. Yu, Text matching improves sequential recommendation by reducing popularity biases, in: CIKM, 2023, pp. 1534–1544.
[111] C. He, Y. Liu, Q. Li, W. Wang, X. Fu, X. Fu, C. Hong, X. Yao, Multi-grained preference enhanced transformer for multi-behavior sequential recommendation, in: SIGKDD, 2025, pp. 872–883.
[112] L. Liu, L. Cai, C. Zhang, X. Zhao, J. Gao, W. Wang, Y. Lv, W. Fan, Y. Wang, M. He, Z. Liu, Q. Li, LinRec: Linear attention mechanism for long-term sequential recommender systems, in: SIGIR, 2023, pp. 289–299.
[113] P. Zhang, J. Guo, C. Li, Y. Xie, J. Kim, Y. Zhang, X. Xie, H. Wang, S. Kim, Efficiently leveraging multi-level user intent for session-based recommendation via atten-mixer network, in: WSDM, 2023, pp. 168–176.
[114] Y. Xie, P. Zhou, S. Kim, Decoupled side information fusion for sequential recommendation, in: SIGIR, 2022, pp. 1611–1621.
[115] Y. Dong, N.V. Chawla, A. Swami, Metapath2vec: Scalable representation learning for heterogeneous networks, in: SIGKDD, 2017, pp. 135–144.
[116] T. Dao, A. Gu, Transformers are SSMs: Generalized models and efficient algorithms through structured state space duality, in: ICML, 2024, pp. 1–52.
[117] Z. Wan, X. Liu, B. Wang, J. Qiu, B. Li, T. Guo, G. Chen, Y. Wang, Spatio-temporal contrastive learning-enhanced GNNs for session-based recommendation, ACM Trans. Inf. Syst. 42 (2) (2024) 58:1–58:26.
[118] J. Li, X. Wang, Z. Tu, M.R. Lyu, On the diversity of multi-head attention, Neurocomputing 454 (2021) 14–24.

Weiyue Li received the Ph.D. degree in Management Science and Engineering (E-Commerce) from School of Management Science and Engineering, Dongbei University of Finance and Economics, in 2026. His research interests include recommendation systems, time series analysis, and social network analysis. His publications have appeared in DSS, IP&M, ESWA, EAAI, among others.

Ming Gao received the BE, MS and Ph.D. degree in information technology and management from the School of Management Science and Engineering (SMSE), Dongbei University of Finance and Economics (DUFE), China, in 2002, 2004 and 2013. He is currently a professor and the head of department of big data management and application with SMSE, DUFE. He is also the deputy director of Key Laboratory of Big Data Management Optimization and Decision of Liaoning Province. His research interests include cloud computing, deep learning, and big data science and applications. He has published over 60 peer-reviewed journal and conference papers such as in IEEE Trans, ACM TOIS, JASIST, DSS, Advanced Science, Bioinformatics, IP&M, AOR, BIB and AAAI.

Jingmin An received the Ph.D. degree in computer science from School of Information Science and Technology, Dalian Maritime university, in 2024. He is currently an associate professor in management science and engineering with Dongbei University of Finance and Economics, China. His major research interests include recommender system, ubiquitous computing, and knowledge engineering. His publications have appeared in ACM TOIS, Annals of Operations Research, DSS, among others.

Bowei Chen is a Professor of Business Analytics and Artificial Intelligence at the Adam Smith Business School, University of Glasgow, UK. He received his Ph.D. in Computer Science from University College London, and his research lies at the intersection of artificial intelligence, machine learning, and business analytics, focusing on probabilistic methods, large-scale analytical frameworks, and responsible AI for business applications. His work has been published in venues including IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Network Science and Engineering, ACM Transactions on Knowledge Discovery from Data, ACM Transactions on Intelligent Systems and Technology, European Journal of Operational Research, Risk Analysis, Tourism Management, SIGKDD, SIGIR, and AAAI.

Jiafu Tang received the M.Sc. degree in automation and the Ph.D. degree in system engineering, both from the Northeastern University, Shenyang, China. He is currently the Changjiang Scholarship Chair Professor and once served as Dean of the College of Management Science and Engineering, Dongbei University of Finance and Economics. He has been selected as a Highly Cited Chinese Researcher by Elsevier in the discipline of Management Science and Engineering for five consecutive years (2020–2024). He is the author or coauthor of more than 100 papers published in international and local journals, among which 60 are in refereed international journals. His research interests include operations management and optimization, supply chain planning and logistics management, quality design in new product development, and quality function deployment.

Xuhui Wang is a professor , doctoral supervisor, and the President of Dongbei University of Finance and Economics, China. He is also the Vice President of the Chinese Institute of Business Administration, China Society of Logistics, and China Information Economics Society. His research interests include online consumer behavior and social media. His research has appeared (or is forthcoming) in Annals of Tourism Research, Tourism Management, International Journal of Contemporary Hospitality Management, etc.

Yeming (Yale) Gong is a full professor at EMLYON Business School. He published 150+ articles in journals including MISQ, Production and Operations Management, Transportation Science, and DSS.