SHORT-PAPER

GEAR: Generalized Alternating Regressor for Multi-Behavior Sequential Recommendation

JUNZHE JIANG, University of Science and Technology of China, Hefei, Anhui, China

KAI ZHANG, University of Science and Technology of China, Hefei, Anhui, China

JUNFENG KANG, University of Science and Technology of China, Hefei, Anhui, China

YUCONG LUO, University of Science and Technology of China, Hefei, Anhui, China

MIN GAO, University of Science and Technology of China, Hefei, Anhui, China

SIGIR '25: Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (July 2025)
https://doi.org/10.1145/3726302.3730200
ISBN: 9798400715921

GEAR: Generalized Alternating Regressor for Multi-Behavior Sequential Recommendation

Junzhe Jiang
State Key Laboratory of Cognitive Intelligence, University of Science and Technology of China
Hefei, Anhui, China
jzjiang@mail.ustc.edu.cn

Kai Zhang*
State Key Laboratory of Cognitive Intelligence, University of Science and Technology of China
Hefei, Anhui, China
kkzhang08@ustc.edu.cn

Junfeng Kang
State Key Laboratory of Cognitive Intelligence, University of Science and Technology of China
Hefei, Anhui, China
kangjf@mail.ustc.edu.cn

Yucong Luo
State Key Laboratory of Cognitive Intelligence, University of Science and Technology of China
Hefei, Anhui, China
prime666@mail.ustc.edu.cn

Min Gao
The First Affiliated Hospital of the USTC, University of Science and Technology of China
Hefei, Anhui, China
gm197866@mail.ustc.edu.cn

Abstract

Modern recommender systems face a critical challenge in modeling the intricate interplay between multi-behavior interactions of users (e.g., clicks, adds-to-cart and purchases) and temporal dynamics that drive evolving preferences. While existing multi-behavior sequential recommendation methods attempt to capture these signals, they often suffer from fragmented modeling, such as decoupling behaviors and items into separate sequences, neglecting time-aware transitions, or relying on computationally intensive architectures that hinder real-world scalability. To address these limitations, we propose GEneralized Alternating Regressor (GEAR), a novel framework that unifies behaviors, items, and temporal contexts into a single autoregressive sequence through an alternating architecture. At its core, GEAR represents user interactions as triplets and processes them through a modular transformer architecture. In this architecture, each triplet is alternately modeled at lower layers to disentangle fine-grained patterns, while upper layers jointly learn cross-signal dependencies. This design mimics the interlocking mechanism of gears, enabling the seamless transitions between multi-behavior dynamics and item transitions. Additionally, we incorporate a time-bias term to quantify the decay of behavioral influence across both short- and long-term horizons. Extensive experiments on real-world datasets validate the effectiveness, generalizability, and computational efficiency of the proposed framework.

CCS Concepts

• **Information systems → Recommender systems**.

Keywords

Multi-Behavior Sequential Recommendation, Autoregressive Sequence Modeling, Temporal Dynamics

ACM Reference Format:
Junzhe Jiang, Kai Zhang, Junfeng Kang, Yucong Luo, and Min Gao. 2025. GEAR: Generalized Alternating Regressor for Multi-Behavior Sequential Recommendation. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '25), July 13–18, 2025, Padua, Italy*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3726302.3730200

1 INTRODUCTION

Recommender systems have demonstrated immense value across diverse domains such as e-commerce, social media, and online education. By analyzing users' historical interactions, these systems [1, 6, 27, 29, 34] accurately capture user preferences and deliver personalized services. However, traditional systems often focus solely on modeling single-behavior sequences, overlooking the latent dependencies between multiple user behaviors, such as clicking, adding to a cart, and purchasing [13, 28, 32, 33, 35]. This limitation constrains further improvements in recommendation quality.

Multi-behavior sequential recommendation (MBSR) combines behavior and item sequences, offering an integrated perspective on the dynamic changes in user preferences and behavioral patterns. Early efforts [2, 7, 11, 22] explored modeling different behavior correlations through multi-task learning, graph neural networks, and meta-learning but treated behaviors as static, overlooking temporal evolution. Subsequent works [3, 4, 19] adopted a two-stage paradigm, first generating subsequences for each behavior and then aggregating them through weighted summation or attention mechanisms. More recent methods [10, 24, 25, 36] have sought to enhance performance by treating auxiliary behavior sequences and primary item sequences as separate inputs and modeling their dependencies.

Despite their advancements, existing methods introduce considerable complexity and limitations [12, 15, 26, 30]. Two-stage approaches struggle to handle scenarios where users exhibit multiple behaviors on the same item due to decoupled behavior modeling.

*Corresponding author

SIGIR '25, Padua, Italy.

ACM ISBN 979-8-4007-1592-1/25/07
https://doi.org/10.1145/3726302.3730200

Meanwhile, methods that model both auxiliary and primary sequences often build complex transfer models to iteratively extract information from the auxiliary sequences. Such architectures introduce complexity through cascaded modules and face challenges in incremental updates, hindering real-world deployment.

To address these challenges, we propose the GEneralized Alternating Regressor (GEAR) for multi-behavior sequential recommendation. Our model adopts a time-centric perspective, alternating between behaviors and items within a unified framework and training them autoregressively. Specifically, we model each user interaction as a triplet, alternately predicting the user's next behavior and the next item they are likely to engage with, based on the current behavior and time gap. This alternating process resembles the interlocking mechanism of gears, where each behavior or item drives the next, resulting in smooth and continuous motion. Moreover, the architecture separates behaviors and items at lower levels, while jointly modeling them at higher levels to bridge the semantic gap. Additionally, a time difference bias term is incorporated to better capture both short-term and long-term preferences. Overall, GEAR reduces parameter complexity and improves efficiency, providing a flexible solution for various recommendation tasks.

The key contributions of this work are as follows:

- We introduce a unified framework to jointly model behavior, time, and item sequences, bridging the semantic gap between user behaviors and items.
- We design an efficient structure that reduces complexity while being easily expandable. We also explicitly introduce temporal bias to more accurately capture differences in user interests in the long and short term.
- Through extensive experiments, we demonstrate the superior performance of our model across multiple datasets, showcasing its advantages in quality and efficiency.

2 PROBLEM FORMULATION

To formalize the problem, we introduce the following notations. Let $U = \{u_1, u_2, \ldots, u_{|U|}\}$ and $V = \{v_1, v_2, \ldots, v_{|V|}\}$ represent the set of users and items, respectively, where $|U|$ and $|V|$ denote the number of users and items. The set of behaviors is denoted as $B = \{b_1, b_2, \ldots, b_{|B|}\}$, where $|B|$ represents the number of behavior types. Among these, one type of behavior, such as "purchase," is considered the target behavior, while the remaining types are regarded as auxiliary behaviors.

For each user $u \in U$, their historical interactions can be represented as a sequence: $S_u = \{(v_u^1, b_u^1, t_u^1), \ldots, (v_u^L, b_u^L, t_u^L)\}$, where (v_u^i, b_u^i, t_u^i) denotes the i-th interaction triplet, describing that item $v_u^i \in V$ was interacted with by user u under behavior $b_u^i \in B$ at timestamp $t_u^i \in \mathbb{R}$, and L is the sequence length. The interactions are ordered chronologically.

3 METHODOLOGY

3.1 Time-Grounded Alternating Encoding

Given a user's historical interaction sequence $S_u = \{(v_u^i, b_u^i, t_u^i)\}_{i=1}^L$, where v_u^i, b_u^i, and t_u^i represent the i-th item, behavior, and timestamp, respectively, we construct an alternating sequence $\mathcal{A}_u$ that interleaves behavior and item tokens:

$$\mathcal{A}_u = [\psi_b(b_u^1), \psi_v(v_u^1), \ldots, \psi_b(b_u^L), \psi_v(v_u^L)], \tag{1}$$

where ψ_b and ψ_v represent learnable embeddings layer for behaviors and items, respectively. Figure 1 shows the model architecture of our proposed GEAR.

3.2 Modular Attention with Temporal Bias

To disentangle behavior-item dependencies while capturing temporal dynamics, GEAR introduces temporal influence decay and uses two specialized attention modules, which we will describe in turn.

Temporal Influence Decay: Time significantly affects user decision-making, with preferences varying across short-term and long-term intervals [9, 31]. Users' purchasing behaviors are also influenced by specific periods like seasons and holidays. To quantify how behavioral impacts diminish over time, we introduce a time-bias term $\Delta_{ij} = \log(t_u^j - t_u^i)$ for $j > i$, which captures the decay of behavioral influence over time. To address the challenges in modeling raw timestamps and linear time gaps, we propose a temporal decay bias function $\Phi(\Delta t) = -\alpha \cdot \log(1 + \Delta t)$, where $\alpha \in \mathbb{R}^{n_h}$ denotes the learnable parameters for each attention head [14]. This continuous formulation effectively captures temporal relationships and adapts to downweight the influence of older interactions.

Behavior-Specific Attention: Given the behavior tokens $\mathcal{B} = \{\mathcal{A}_u[2k]\}$, the behavior-specific attention is computed as:

$$\mathrm{Attn}_b(\mathcal{B}) = \mathrm{Softmax}\left(\frac{Q_b K_b^\top}{\sqrt{d}} + \Phi(\Delta t)\right) V_b, \tag{2}$$

where Q_b, K_b, and V_b are the query, key, and value matrices for behavior tokens, and d is the dimension of these matrices.

Item-Specific Attention: For the item tokens $\mathcal{I} = \{\mathcal{A}_u[2k+1]\}$, the item-specific attention is calculated as:

$$\mathrm{Attn}_v(\mathcal{I}) = \mathrm{Softmax}\left(\frac{Q_v K_v^\top}{\sqrt{d}}\right) V_v, \tag{3}$$

where Q_v, K_v, and V_v represent the query, key, and value matrices for item tokens. By separating the processing of behaviors and items in lower layers, GEAR significantly reduces the computational complexity from $O(L^2 d)$ to $O((L^2/2)d)$.

3.3 Hierarchical Cross-Signal Fusion

The alternating tokens undergo multiple layers of processing to effectively capture and fuse cross-signal dependencies.

Low-Level Disentanglement: Behavior and item attention blocks are applied to each modality separately, allowing for independent processing of behavior and item sequences. The outputs of the behavior and item sequences are as follows:

$$\mathcal{B}' = \mathrm{LayerNorm}(\mathcal{B} + \mathrm{Attn}_b(\mathcal{B})), \tag{4}$$

$$\mathcal{I}' = \mathrm{LayerNorm}(\mathcal{I} + \mathrm{Attn}_v(\mathcal{I})). \tag{5}$$

High-Level Fusion: The processed behavior and item representations, $\mathcal{B}'$ and $\mathcal{I}'$ are concatenated into a unified sequence $\mathcal{H}_0 = [\mathcal{B}', \mathcal{I}']$, which is then passed through a multi-layer transformer to learn high-level cross-signal dependencies:

$$\mathcal{H}_{l+1} = \mathrm{Transformer}(\mathcal{H}_l), \quad l = 0, \ldots, N-1. \tag{6}$$

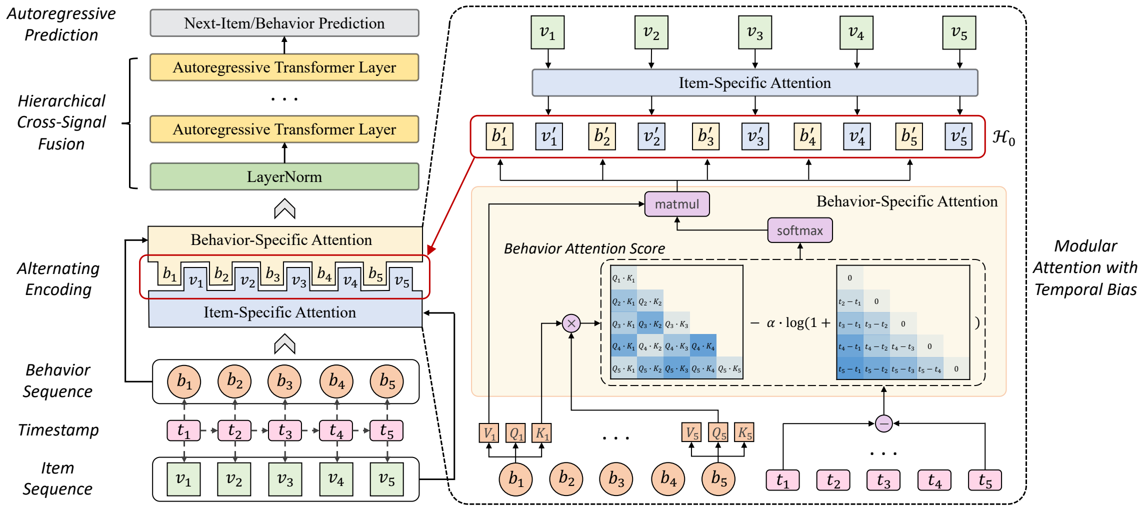

Figure 1: The model architecture of GEAR. The zoom-in view of the modular attention with temporal bias is in the right panel.

Table 1: Statistics of the preprocessed datasets. "Avg. L" denotes the average length of item sequences. "Avg. i" denotes the average number of interactions associated with an item.

Dataset	# Users	# Items	# Interactions	Avg. L	Avg. i
Retail	147,894	99,037	7,658,926	51.66	77.13
IJCAI	423,423	874,328	36,222,123	84.91	41.12
Yelp	19,800	22,734	1,400,002	70.22	61.15

3.4 Autoregressive Prediction

The final hidden states $\mathcal{H}_N$ are split into separate behavior and item streams to facilitate prediction. The next behavior b_u^{L+1} is predicted from the penultimate behavior token embedding $\mathcal{H}_N^{[-2]}$ using a softmax function:

$$P(b^{L+1}|S_u) = \text{Softmax}(\mathcal{H}_N^{[-2]}\psi_b^\top), \tag{7}$$

The next item v_u^{L+1} is predicted from the last item token embedding $\mathcal{H}_N^{[-1]}$, conditioned on the predicted behavior b_u^{L+1}:

$$P(v^{L+1}|S_u, b^{L+1}) = \text{Softmax}(\mathcal{H}_N^{[-1]}\psi_v^\top), \tag{8}$$

where $\psi_b^\top \in \mathbb{R}^{|B|\times d}$ and $\psi_v^\top \in \mathbb{R}^{|V|\times d}$ are the behavior embedding matrix and item embedding matrix in Section 3.1.

To optimize the model, we minimize the joint negative log-likelihood, which combines both behavior and item prediction tasks:

$$\mathcal{L} = -\sum_u \left(\lambda \log P(b_u^{L+1}|S_u) + \log P(v_u^{L+1}|S_u, b_u^{L+1})\right), \tag{9}$$

where λ is a hyperparameter that controls the balance between behavior and item prediction.

4 EXPERIMENTS

4.1 Experimental Setting

4.1.1 Datasets. We evaluate GEAR on three multi-behavior recommendation benchmarks: **Retail**, **IJCAI**, and **Yelp** [21, 22, 25]. **Retail** is sourced from Taobao e-commerce logs and includes four interaction types: *page view*, *tag-as-favorite*, *add-to-cart*, and *purchase*. **IJCAI** originates from an IJCAI competition and shares the same behavior types as Retail, but is larger and sparser. **Yelp** is derived from public review data, where ratings are transformed into three behaviors: *dislike* (rating ≤ 2), *neutral* (2 < rating < 4), and *like* (rating ≥ 4). We preprocess the dataset by filtering out items that appear fewer than 5 times to ensure data quality [21, 25]. Table 1 summarizes the key statistics of the datasets.

4.1.2 Baselines. We compare the proposed approach with the following baseline methods. **Traditional recommendation** [20, 23] focus on learning user and item representations through matrix factorization or graph-based techniques. **Sequential recommendation** [5, 8, 16, 17] capture sequential dependencies in user behavior, often using recurrent or transformer-based architectures. **Multi-behavior recommendation** [2, 7, 22] incorporate multiple types of user behaviors by considering inter-behavior dependencies. **Multi-behavior sequential recommendation** [3, 4, 10, 12, 15, 18, 19, 24, 25] focus on modeling both behavior-specific and multi-behavior dependencies in a sequential manner.

4.1.3 Implementation details. We implement the proposed model using PyTorch. We set the training batch size as 128 and max sequence length L as 50 [15, 25]. The embedding dimension is set to 16 for the Yelp dataset with number of transformer layers $N = 2$, while the Retail and IJCAI datasets use 32 embedding dimensions alongside $N = 4$. For the baselines with transformer architecture, we maintain the same settings as above. The hyperparameter λ is set to 0.1. We make our codes publicly available [1].

4.2 Overall Performance

We evaluate the performance of our proposed GEAR on three benchmark datasets, which have been extensively employed in prior studies, ensuring a fair comparison with state-of-the-art methods. Table 2 demonstrates that our proposed method surpasses the established baseline and validates the effectiveness of GEAR in capturing complex multi-behavior patterns and temporal dynamics. The autoregressive structure of GEAR allows for online updates via sliding window processing, making it well-suited for real-world recommendation systems that require frequent model updates based on new user interactions.

[1] https://github.com/Gnimixy/GEAR/

Table 2: Experimental results on three datasets. The best results are boldfaced and the second-best results are underlined.

	Retail				IJCAI				Yelp			
	HR@5	NDCG@5	HR@10	NDCG@10	HR@5	NDCG@5	HR@10	NDCG@10	HR@5	NDCG@5	HR@10	NDCG@10
DMF [23]	0.189	0.135	0.298	0.185	0.275	0.197	0.379	0.242	0.655	0.415	0.743	0.474
NGCF [20]	0.196	0.130	0.318	0.197	0.346	0.232	0.463	0.293	0.674	0.428	0.778	0.495
GRU4Rec [5]	0.241	0.176	0.368	0.215	0.498	0.347	0.592	0.391	0.684	0.432	0.781	0.495
Caser [17]	0.204	0.132	0.341	0.193	0.329	0.213	0.439	0.274	0.673	0.425	0.763	0.489
SASRec [8]	0.253	0.182	0.377	0.226	0.489	0.351	0.604	0.412	0.728	0.451	0.793	0.501
BERT4Rec [16]	0.278	0.193	0.396	0.241	0.493	0.362	0.622	0.426	0.745	0.473	0.824	0.528
NMTR [2]	0.183	0.159	0.328	0.168	0.396	0.269	0.495	0.312	0.694	0.445	0.785	0.467
MB-GCN [7]	0.247	0.172	0.362	0.215	0.345	0.209	0.453	0.269	0.734	0.468	0.807	0.513
MB-GMN [22]	0.394	0.221	0.498	0.304	0.467	0.286	0.524	0.337	0.736	0.525	0.862	0.580
DIPN [4]	0.256	0.157	0.332	0.184	0.425	0.259	0.481	0.305	0.557	0.452	0.683	0.517
BINN [10]	0.358	0.224	0.463	0.306	0.398	0.242	0.469	0.291	0.375	0.344	0.489	0.413
MGNN-SPred [19]	0.353	0.218	0.459	0.304	0.381	0.234	0.443	0.279	0.372	0.335	0.485	0.402
DMT [3]	0.547	0.355	0.654	0.408	0.592	0.354	0.675	0.458	0.549	0.471	0.662	0.525
ASLI [18]	0.382	0.255	0.503	0.325	0.436	0.278	0.502	0.314	0.417	0.392	0.528	0.457
MBGen [12]	0.486	0.380	0.533	0.395	0.490	0.350	0.629	0.395	0.398	0.376	0.492	0.431
MBHT [24]	0.687	0.590	0.764	0.615	0.774	0.675	0.852	0.701	0.727	0.550	0.861	0.594
MB-STR [25]	0.691	0.585	0.772	0.611	0.805	0.692	0.883	0.717	0.750	0.574	0.874	0.615
PBAT [15]	0.712	0.630	0.794	0.656	0.848	0.743	0.897	0.761	0.741	0.575	0.867	0.619
GEAR	**0.776**	**0.701**	**0.834**	**0.720**	**0.904**	**0.818**	**0.945**	**0.833**	**0.769**	**0.591**	**0.887**	**0.628**
Impv.	8.99%	11.3%	5.03%	9.76%	6.60%	10.1%	5.35%	9.46%	2.53%	2.78%	1.49%	1.45%

Table 3: Ablation study on key components of GEAR. Note that the cut-off of the ranked list is 10.

	Retail		IJCAI		Yelp	
	HR	NDCG	HR	NDCG	HR	NDCG
*w/o **AE***	0.377	0.226	0.604	0.412	0.793	0.501
*w/o **BP***	0.625	0.377	0.653	0.441	0.824	0.558
*w/o **LLF***	0.782	0.649	0.865	0.885	0.862	0.596
*w/o **TID***	0.805	0.677	0.921	0.795	0.873	0.616
GEAR	**0.834**	**0.720**	**0.945**	**0.833**	**0.887**	**0.628**

4.3 Ablation Study

To validate the rationale behind GEAR's design, we conduct ablation studies on four variants: (1) **w/o TID** removes the Temporal Influence Decay mechanism; (2) **w/o LLF** eliminates Low-Level Disentanglement; (3) **w/o BP** disables Behavior Prediction by setting $\lambda = 0$ in function (9); (4) **w/o AE** abandons Alternating Encoding. As shown in Table 3, combining alternating encoding with low-level disentanglement and temporal bias yields more accurate modeling of both short- and long-term user preferences.

4.4 Efficiency Analysis

Figure 2(a) illustrates the trade-off between model size, training efficiency, and performance across state-of-the-art MBSR methods. Experiments were conducted with identical settings, plotting parameter counts against average epoch duration. Our analysis reveals that while baseline models improve performance at the cost of increased parameters and training time, GEAR achieves superior accuracy with 73% fewer parameters and 3.8× faster training than the second-best method PBAT. This efficiency stems from the alternating encoder eliminates redundant cross-module connections.

4.5 Impact of Auxiliary Behaviors

We conduct additional ablation studies to evaluate the impact of different auxiliary behaviors in our model. Specifically, we explore four variants on the IJCAI dataset: "-pv", "-fav", and "-cart" represent the GEAR model without incorporating page view, tag-as-favorite, and add-to-cart behaviors, respectively, while "+buy" refers to the variant that solely utilizes the target behavior (purchase). The results in figure 2(b) show that removing any auxiliary behavior leads to performance degradation, with the most significant decline observed when the page view behavior is excluded.

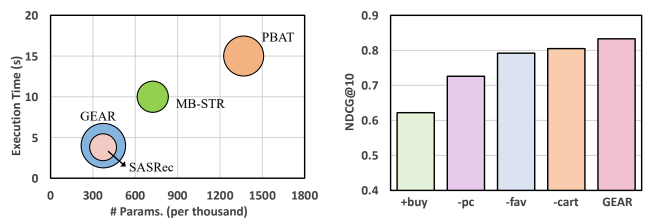

Figure 2: (a) Left: Efficiency comparison of baseline models. Circle centers denote parameter counts (x-axis, thousands) vs. training time per epoch (y-axis, seconds). Circle radii represent relative NDCG@10 performance. (b) Right: Impact of auxiliary behavior removal on recommendation accuracy.

5 CONCLUSION

In this paper, we introduce GEAR, a novel framework for MBSR, which addresses the challenges of modeling the complex user behavior and temporal dynamics. Our approach employs an innovative alternating encoding strategy to disentangle fine-grained patterns at lower layers and capture cross-signal dependencies at higher layers. Moreover, integrating a time difference bias term enables the model to effectively distinguish between short-term urgency and long-term preferences. Extensive experiments on three real-world datasets demonstrate that GEAR consistently outperforms state-of-the-art baselines in terms of accuracy and efficiency.

Acknowledgments

This research was partially supported by the National Natural Science Foundation of China (Grants No.62406303), Anhui Provincial Natural Science Foundation (No. 2308085QF229), Anhui Province Science and Technology Innovation Project (202423k09020010) and the Fundamental Research Funds for the Central Universities (No. WK2150110034).

References

[1] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*. 191–198.

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

[3] Yulong Gu, Zhuoye Ding, Shuaiqiang Wang, Lixin Zou, Yiding Liu, and Dawei Yin. 2020. Deep multifaceted transformers for multi-objective ranking in large-scale e-commerce recommender systems. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2493–2500.

[4] Long Guo, Lifeng Hua, Rongfei Jia, Binqiang Zhao, Xiaobo Wang, and Bin Cui. 2019. Buying or browsing?: Predicting real-time purchasing intent using attention-based deep network with multiple behavior. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1984–1992.

[5] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2015. Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939* (2015).

[6] Junzhe Jiang, Shang Qu, Mingyue Cheng, Qi Liu, Zhiding Liu, Hao Zhang, Rujiao Zhang, Kai Zhang, Rui Li, Jiatong Li, and Min Gao. 2024. Reformulating Sequential Recommendation: Learning Dynamic User Interest with Content-enriched Language Modeling. In *Database Systems for Advanced Applications: 29th International Conference, DASFAA 2024, Gifu, Japan, July 2–5, 2024, Proceedings, Part III*. Springer-Verlag, Berlin, Heidelberg, 353–362. doi:10.1007/978-981-97-5555-4_25

[7] Bowen Jin, Chen Gao, Xiangnan He, Depeng Jin, and Yong Li. 2020. Multi-behavior recommendation with graph convolutional networks. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 659–668.

[8] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *2018 IEEE Int'l Conf. on Data Mining (ICDM)*. IEEE, 197–206.

[9] Jiacheng Li, Yujie Wang, and Julian McAuley. 2020. Time interval aware self-attention for sequential recommendation. In *Proceedings of the 13th international conference on web search and data mining*. 322–330.

[10] Zhi Li, Hongke Zhao, Qi Liu, Zhenya Huang, Tao Mei, and Enhong Chen. 2018. Learning from history and present: Next-item recommendation via discriminatively exploiting user behaviors. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1734–1743.

[11] Yang Liu, Qianzhen Rao, Weike Pan, and Zhong Ming. 2023. Variational Collective Graph AutoEncoder for Multi-behavior Recommendation. In *2023 IEEE International Conference on Data Mining (ICDM)*. IEEE, 438–447.

[12] Zihan Liu, Yupeng Hou, and Julian McAuley. 2024. Multi-behavior generative recommendation. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. 1575–1585.

[13] Qi Pi, Guorui Zhou, Yujing Zhang, Zhe Wang, Lejian Ren, Ying Fan, Xiaoqiang Zhu, and Kun Gai. 2020. Search-based user interest modeling with lifelong sequential behavior data for click-through rate prediction. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 2685–2692.

[14] Ofir Press, Noah Smith, and Mike Lewis. 2021. Train Short, Test Long: Attention with Linear Biases Enables Input Length Extrapolation. In *International Conference on Learning Representations*.

[15] Jiajie Su, Chaochao Chen, Zibin Lin, Xi Li, Weiming Liu, and Xiaolin Zheng. 2023. Personalized behavior-aware transformer for multi-behavior sequential recommendation. In *Proceedings of the 31st ACM International Conference on Multimedia*. 6321–6331.

[16] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. 2019. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM international conference on information and knowledge management*. 1441–1450.

[17] Jiaxi Tang and Ke Wang. 2018. Personalized top-n sequential recommendation via convolutional sequence embedding. In *Proceedings of the eleventh ACM international conference on web search and data mining*. 565–573.

[18] Md Mehrab Tanjim, Congzhe Su, Ethan Benjamin, Diane Hu, Liangjie Hong, and Julian McAuley. 2020. Attentive sequential models of latent intent for next item recommendation. In *Proceedings of The Web Conference 2020*. 2528–2534.

[19] Wen Wang, Wei Zhang, Shukai Liu, Qi Liu, Bo Zhang, Leyu Lin, and Hongyuan Zha. 2020. Beyond clicks: Modeling multi-relational item graph for session-based target behavior prediction. In *Proceedings of The Web Conference 2020*. 3056–3062.

[20] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. 2019. Neural graph collaborative filtering. In *Proceedings of the 42nd international ACM SIGIR conference on Research and development in Information Retrieval*. 165–174.

[21] Lianghao Xia, Chao Huang, Yong Xu, Peng Dai, Bo Zhang, and Liefeng Bo. 2020. Multiplex behavioral relation learning for recommendation via memory augmented transformer network. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2397–2406.

[22] Lianghao Xia, Yong Xu, Chao Huang, Peng Dai, and Liefeng Bo. 2021. Graph meta network for multi-behavior recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 757–766.

[23] Hong-Jian Xue, Xinyu Dai, Jianbing Zhang, Shujian Huang, and Jiajun Chen. 2017. Deep Matrix Factorization Models for Recommender Systems.. In *IJCAI*, Vol. 17. Melbourne, Australia, 3203–3209.

[24] Yuhao Yang, Chao Huang, Lianghao Xia, Yuxuan Liang, Yanwei Yu, and Chenliang Li. 2022. Multi-behavior hypergraph-enhanced transformer for sequential recommendation. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2263–2274.

[25] Enming Yuan, Wei Guo, Zhicheng He, Huifeng Guo, Chengkai Liu, and Ruiming Tang. 2022. Multi-Behavior Sequential Transformer Recommender. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1642–1652.

[26] Jiaqi Zhai, Lucy Liao, Xing Liu, Yueming Wang, Rui Li, Xuan Cao, Leon Gao, Zhaojie Gong, Fangda Gu, Michael He, et al. 2024. Actions speak louder than words: Trillion-parameter sequential transducers for generative recommendations. *arXiv preprint arXiv:2402.17152* (2024).

[27] Hao Zhang, Mingyue Cheng, Qi Liu, Zhiding Liu, Junzhe Jiang, and Enhong Chen. 2025. AutoSAM: Towards Automatic Sampling of User Behaviors for Sequential Recommender Systems. doi:10.48550/arXiv.2311.00388 arXiv:2311.00388 [cs]

[28] Kai Zhang, Qi Liu, Hao Qian, Biao Xiang, Qing Cui, Jun Zhou, and Enhong Chen. 2021. Eatn: An efficient adaptive transfer network for aspect-level sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering* 35, 1 (2021), 377–389.

[29] Kai Zhang, Hao Qian, Qing Cui, Qi Liu, Longfei Li, Jun Zhou, Jianhui Ma, and Enhong Chen. 2021. Multi-interactive attention network for fine-grained feature learning in ctr prediction. In *Proceedings of the 14th ACM international conference on web search and data mining*. 984–992.

[30] Kai Zhang, Hefu Zhang, Qi Liu, Hongke Zhao, Hengshu Zhu, and Enhong Chen. 2019. Interactive attention transfer network for cross-domain sentiment classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 5773–5780.

[31] Rujiao Zhang, Hao Zhang, Yucong Luo, Zhiding Liu, Mingyue Cheng, Qi Liu, and Enhong Chen. 2024. Learning the Dynamics in Sequential Recommendation by Exploiting Real-time Information. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*. 4288–4292.

[32] Tingting Zhang, Pengpeng Zhao, Yanchi Liu, Victor S Sheng, Jiajie Xu, Deqing Wang, Guanfeng Liu, Xiaofang Zhou, et al. 2019. Feature-level Deeper Self-Attention Network for Sequential Recommendation.. In *IJCAI*. 4320–4326.

[33] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 1059–1068.

[34] Haolin Zhou, Junwei Pan, Xinyi Zhou, Xihua Chen, Jie Jiang, Xiaofeng Gao, and Guihai Chen. 2024. Temporal Interest Network for User Response Prediction. In *Companion Proceedings of the ACM on Web Conference 2024*. 413–422.

[35] Kun Zhou, Hui Wang, Wayne Xin Zhao, Yutao Zhu, Sirui Wang, Fuzheng Zhang, Zhongyuan Wang, and Ji-Rong Wen. 2020. S3-rec: Self-supervised learning for sequential recommendation with mutual information maximization. In *Proceedings of the 29th ACM international conference on information & knowledge management*. 1893–1902.

[36] Meizi Zhou, Zhuoye Ding, Jiliang Tang, and Dawei Yin. 2018. Micro behaviors: A new perspective in e-commerce recommender systems. In *Proceedings of the eleventh ACM international conference on web search and data mining*. 727–735.