Information Systems 125 (2024) 102427

Contents lists available at ScienceDirect

Information Systems

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

A survey of sequential recommendation systems: Techniques, evaluation, and future directions

Tesfaye Fenta Boka *, Zhendong Niu, Rama Bastola Neupane

School of Computer Science and Technology, Beijing Institute of Technology, Beijing, 100081, China

ARTICLE INFO

Keywords:
Deep learning
Sequential recommendation
Evaluation
Recommender systems

ABSTRACT

Recommender systems are powerful tools that successfully apply data mining and machine learning techniques. Traditionally, these systems focused on predicting a single interaction, such as a rating between a user and an item. However, this approach overlooks the complexity of user interactions, which often involve multiple interactions over time, such as browsing, adding items to a cart, and more. Recent research has shifted towards leveraging this richer data to build more detailed user profiles and uncover complex user behavior patterns. Sequential recommendation systems have gained significant attention recently due to their ability to model users' evolving preferences over time. This survey explores how these systems utilize interaction history to make more accurate and personalized recommendations. We provide an overview of the techniques employed in sequential recommendation systems, discuss evaluation methodologies, and highlight future research directions. We categorize existing approaches based on their underlying principles and evaluate their effectiveness in various application domains. Additionally, we outline the challenges and opportunities in sequential recommendation systems.

Contents

* Corresponding author.
E-mail addresses: tfenta@bit.edu.cn (T.F. Boka), zniu@bit.edu.cn (Z. Niu), rama@bit.edu.cn (R.B. Neupane).

https://doi.org/10.1016/j.is.2024.102427
Received 17 June 2024; Received in revised form 10 July 2024; Accepted 10 July 2024
Available online 16 July 2024

1. Introduction

With the proliferation of the Internet, information overload has become an increasing problem in people's everyday lives [1,2]. Recommender systems effectively alleviate this issue by helping users find desired information and increasing service providers' traffic and revenue [3]. These systems are widely used in various applications such as e-commerce [4], social media sites, news portals, app stores, and digital libraries. They have become one of the most ubiquitous user-centered artificial intelligence applications in modern information systems. Furthermore, recommender systems are a crucial part of artificial intelligence (AI) [5], helping users uncover hidden interests. Recommender systems have evolved significantly since the 1990s [1], starting with basic heuristics for content-based and collaborative filtering. In the late 2000s, matrix factorization emerged as a dominant approach, but its limitations in handling complex user interactions and intricate item features became apparent. The mid-2010s witnessed a revolution in machine learning with the rise of deep learning [1], whose success in areas like speech recognition and natural language processing sparked a new wave of research. Recognizing the potential of deep learning to handle complex data patterns, researchers began applying these techniques to recommender systems, leading to significant advancements in recent years. Recommender systems are crucial for navigating today's information overload. Traditional approaches, such as collaborative and content-based filtering, are effective in many scenarios. However, they cannot account for the dynamic nature of user interests and the sequential nature of user interactions with items. Recommender Systems (RS) are software applications that support users in finding items of interest within extensive collections of objects, often in a personalized way [4]. Today, such systems are used in various application domains, including e-commerce and media streaming. Receiving automated recommendations in different forms has become a part of our daily online user experience. Internally, such systems analyze the past behavior of individual users or a user community as a whole to detect patterns in the data. On typical online sites, various relevant user actions can be recorded, such as viewing an item or making a purchase. Several actions by a single user may relate to the same item. These recorded actions and the detected patterns are then used to compute recommendations that match individual users' preference profiles.

In academic environments, the predominant problem abstraction is matrix completion. Here, a user–item rating matrix [6] is given, and the goal is to predict the missing values. This abstraction is generally well-suited to train machine-learning models that aim to capture longer-term user preference profiles. However, these algorithms typically lack specific means to account for users' short-term behavior or intents in their recommendations. Additionally, these algorithms are not designed to utilize the rich information in the sequentially ordered user interaction logs often available in practical applications. Sequential recommendation systems (SRS) address this by considering the order of user interactions [7]. Unlike static models, SRS can analyze a user's history to predict their evolving preferences, leading to more accurate and personalized recommendations. In this survey, we aim to thoroughly examine SRS, covering methodologies, recent advancements, evaluation techniques, and future research directions.

2. Background

In this section, we provide background information on sequential recommendation systems, including key concepts and traditional approaches for comparison.

2.1. Key concepts in SRS

Sequential recommendation systems focus on modeling the sequential behavior of users as they interact with items over time. Key concepts in SRS include:

- **User–Item Interactions**: Representing the historical interactions between users and items, such as clicks, views, purchases, and ratings.
- **Sequence Modeling**: Techniques for capturing the temporal dependencies in user interactions and predicting future behavior based on past interactions.
- **Session-based vs. Long-term Preferences**: Distinguishing between short-term preferences within a session and long-term preferences that evolve over multiple sessions.

2.2. Traditional recommender system approaches

Traditional recommender system approaches serve as a baseline for comparing sequential recommendation techniques. These include:

- **Collaborative Filtering**: Recommending items based on user–item interaction patterns and similarities between users or items. One of the key foundations for recommender systems is collaborative filtering [2]. This approach relies on the idea that users who have shown similar preferences for items in the past are likely to have similar interests in the future.
- **Content-based Filtering**: Recommending items based on their attributes and matching them with users' preferences.

Unique Contributions of the Survey Paper: Comprehensive Taxonomy and Classification: The survey paper offers an in-depth taxonomy and classification of sequential recommendation systems (SRS), categorizing them into traditional approaches and neural network-based approaches. It provides: A clear framework for understanding the evolution and advancements in SRS. Highlighting essential techniques such as Matrix Factorization. Markov Chains. Graph Neural Networks (GNNs). Recurrent Neural Networks (RNNs). Evaluation Metrics and Methodologies:The paper discusses the evaluation methodologies and metrics commonly used in SRS research, such as NDCG, MRR, and HitRate. It emphasizes the importance of accurate and reliable metrics to assess the effectiveness of SRS, addressing issues related to offline evaluation and the need for new datasets with accurate temporal information. Identification of Challenges and Future Directions:The survey identifies significant challenges in SRS, including data sparsity, scalability, and the need for models that capture long-term dependencies. It outlines future research directions, such as context-aware and social-aware sequential recommendation systems, cross-domain recommendation with time-sequential data, and hybrid approaches

Table 1
Sources of selected publications.

S/N	Publication source	Impact factor	Number of selected publications	References
1	IEEE Transactions on Emerging Topics in Computational Intelligence	5.3	1	[7]
2	International Journal of Machine Learning and Cybernetics	5.6	1	[8]
3	Expert Systems With Applications	8.5	5	[9–13]
4	Applied Intelligence	5.3	1	[14]
5	ACM Transactions on Knowledge Discovery from Data	4.4	1	[15]
6	Computer Standards Interfaces	5	1	[16]
7	Neural Computing and Applications	6	1	[17]
8	World Wide Web	3.7	1	[18]
9	IEEE Transactions on Knowledge and Data Engineering	8.9	2	[19,20]
10	IEEE Transactions on Neural Networks and Learning Systems	10.4	2	[21,22]
11	IEEE Transactions on Multimedia	7.3	1	[23]
12	IEEE Transactions on Network Science and Engineering	6.6	1	[24]
13	ACM Transactions on Information Systems	5.6	1	[25]

that integrate various techniques for enhanced accuracy. Novel Insights into Advanced Models: The paper provides novel insights into advanced models used in SRS, mainly focusing on attention-based models, contrastive learning-based models, and graph-based models. It highlights how these models leverage sophisticated mechanisms to capture nuanced user behavior and improve recommendation accuracy, thus pushing the boundaries of traditional recommendation systems.

3. Research methodology

The research methodology employed in this survey paper outlines the systematic approach taken during the survey process. This methodology defines the survey's scope, the search keywords, and the academic archives consulted. The process involved two primary steps. First, we selected papers from top-tier journals and conferences. Second, we analyzed and classified these papers based on their focus on sequential recommendation and relevant methods.

3.1. Selection of publications

The selected published papers were categorized based on the research methodologies employed. Sequential recommendation approaches were thoroughly examined to gain a comprehensive understanding, leading to the formulation of several key research questions:

- **RQ1:** Which type of sequential recommendation system approach performs best?
- **RQ2:** How does the proposed sequential recommendation system learn and represent user behavior?
- **RQ3:** What are the most common evaluation metrics for sequential recommendations?
- **RQ4:** What is the expected future trend of sequential recommendation systems?

3.2. Categorization of the selected publications

Our publication selection process was conducted with meticulous care, focusing specifically on sequential recommendation systems. We targeted approaches and evaluation metrics in these systems, selecting papers from top-tier journals and reputable conferences between 2020 and 2024. Papers were gathered from databases such as IEEE, Expert Systems, and ACM Transactions using search keywords like "recommender+sequential", "sequential recommendation", "session", and "session+recommendation", and applying filtering criteria such as "sequential OR session" and "session AND sequential". This extensive search yielded 100 publications, from which we conducted a detailed examination of sequential recommendations. To ensure the highest relevance of the studies, we implemented a stringent two-step filtering process. The first step involved comprehensively analyzing abstracts, introduction sections, and conclusions. The second step entailed a meticulous review of the proposed methods and the entire articles. The final selection comprised papers that met the following criteria: involvement in sequential or session recommendations and publication in top-tier journals or conferences from 2020 to 2024, with due consideration to the journal's impact factor when applicable. After filtering, 50 publications from 13 sources and their corresponding impact factors were selected and presented in Table 1 for inclusion in this survey. These selected papers form the basis of our comprehensive analysis of sequential recommendation systems.

4. Taxonomy of SRS techniques

This section categorizes and describes different approaches for modeling sequential user behavior in recommendation systems. As shown in Fig. 1, the taxonomy of sequential recommendation techniques is categorized into two main approaches: Traditional Approaches and Neural Network-Based Approaches. These categories encompass various models and techniques that enhance the recommendation process based on user interactions over time.

Traditional approaches. Traditional approaches consist of two primary methods: Matrix Factorization and Markov Chain. Matrix Factorization is a technique that decomposes the user–item interaction matrix into latent factors representing users and items, facilitating the prediction of user preferences. Markov Chain models capture the sequential dependencies between items, predicting the next item based on the previous items in the sequence.

Neural network-based approaches. Neural network-based approaches are divided into Deep Neural Networks and Advanced Models. Deep Neural Networks include several models, such as Graph Convolutional Networks (GCN), Graph Neural Networks (GNN), and Recurrent Neural Networks (RNN). GCN and GNN leverage graph structures to capture complex relationships between items and users, enhancing the recommendation process by considering higher-order connections. RNNs are particularly effective in modeling sequential data due to their ability to maintain hidden states that capture information from previous time steps. Variants of RNNs include Gated Recurrent Units (GRU) and Long Short-Term Memory (LSTM), which address issues such as vanishing gradients and long-term dependencies.

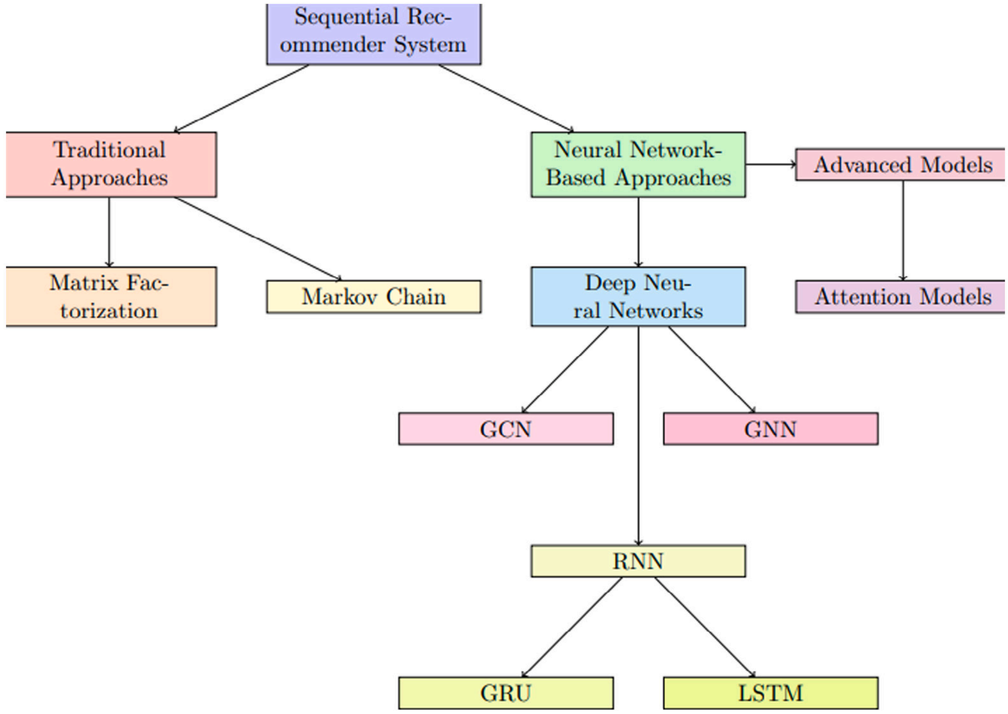

Fig. 1. Taxonomy of sequential recommendation techniques.

Advanced models. Advanced models, categorized under neural network-based approaches, include Attention Models. Attention Models enhance the ability of neural networks to focus on different parts of the input sequence, allowing the model to capture dependencies between distant items more effectively. In general, the taxonomy illustrates a comprehensive framework for sequential recommendation techniques, highlighting the evolution from traditional methods to sophisticated neural network-based models that offer improved capabilities for capturing user preferences and predicting future interactions.

4.1. Markov chain-based models

Markov chain-based recommender systems (RSs) adopt Markov chain models to model transitions over user–item interactions in a sequence to predict the next interaction. According to the specific technique used, Markov chain-based RSs are divided into basic Markov Chain-based approaches and latent Markov embedding-based approaches. The former directly calculates the transition probability based on explicit observations, while the latter embeds the Markov chains into Euclidean space and then calculates the transition probabilities between interactions based on their Euclidean distance [26]. The shortcomings of Markov chain-based RSs are apparent; they can only capture short-term dependencies while ignoring long-term ones due to the Markov property, which assumes that the current interaction depends on one or several most recent interactions only. Additionally, they can capture point-wise dependencies while ignoring collective dependencies over user–item interactions. Consequently, they have been less employed in SRSs in recent years.

4.2. Matrix factorization with sequence modeling

The iALS1 algorithm [27] is designed for learning a matrix factorization model to recommend items to users by analyzing user interactions (implicit feedback) instead of explicit ratings to suggest the top-n items a user might be interested in. Furthermore, the study in [28] introduces a method combining matrix factorization and Markov chains to enhance recommender systems. This method uses personalized transition graphs and factorizes a transition cube with a pairwise interaction model, outperforming traditional matrix factorization and unpersonalized Markov chain models. Fossil [29] combines Matrix Factorization, Markov Chains, and similarity-based methods to predict personalized sequential behavior in recommender systems, showing improved performance on sparse, real-world datasets.

4.3. Deep learning approaches

Deep neural networks can naturally model and capture the comprehensive relationships between entities (e.g., users, items, interactions) in a sequence. Consequently, they have nearly dominated SRSs in the past few years. Recurrent neural networks (RNNs) [26] are the most commonly used deep neural networks for SRSs due to their natural strength in sequence modeling, but they also have defects. Recently, convolutional neural networks (CNNs) and graph neural networks (GNNs) have also been applied in SRSs to compensate for the defects of RNNs. Recurrent Neural Networks (RNNs) have been the backbone of recommender systems, particularly in deep learning approaches. These systems leverage users' past interactions (e.g., purchases or views) to anticipate future actions. RNNs are adept at capturing sequential patterns in this data. However, they are not perfect, and it is essential to be aware of their limitations:

- **Overfitting:** RNNs can sometimes incorrectly assume connections between items that users interact with close together, even if there is no real relationship, leading to inaccurate recommendations.
- **Limited Scope:** RNNs might only focus on predicting the next item based on the most recent interaction, neglecting how past interactions might influence future choices.

Despite these limitations, RNNs with variations like LSTMs and GRUs have been widely used for their ability to model sequential user behavior in recommender systems. Researchers are also exploring new RNN architectures, like hierarchical RNNs, to capture more complex user interaction patterns. Convolutional Neural Networks (CNNs) offer a different approach to RNNs for recommender systems. Instead of processing interactions individually, CNNs treat the entire sequence of user interactions like an "image" with time and hidden factors (latent spaces) as dimensions. They then use filters to analyze this "image" and identify patterns within it. This approach of CNNs brings several advantages over RNNs:

- **No Strict Order Assumption:** CNNs do not make strong assumptions about the order of interactions, making them less prone to overfitting on unrelated items close together in a sequence.
- **Focus on Broader Patterns:** CNNs can identify patterns across the entire sequence, not just between adjacent interactions, potentially capturing more complex user behavior.

However, CNNs also come with their own set of limitations:

- **Limited Long-Term Memory:** The size of the filters used in CNNs restricts their ability to capture long-term dependencies between interactions in a sequence, which can hinder their effectiveness when dealing with user behavior influenced by distant past interactions.

While CNNs offer solutions to some RNN weaknesses, their limitations in capturing long-term dependencies restrict their applications in certain scenarios [26]. Sequential recommendation is crucial in recommending movies and Instagram posts on social media. Deep learning can capture complex user interaction behavior and recommend user preferences and interests based on their past interaction history. Recurrent neural networks perform well in this area, and LSTM and GRU [30] perform well in movies and social media recommendation posts. Caser [31] is a recommendation system that uses recent interactions as a time-aware image to predict what users will interact with next, achieving better results than existing methods. As a study in [7] demonstrates, RP-SANRec is a recommendation model that captures users' evolving interests over time. It combines two techniques: short-term intent learning with GRU and long-term preference modeling with Bi-LSTM. This combination improves the accuracy of user preference prediction and leads to better recommendations. GMRec [10] is a new method for recommending items in a sequence that addresses limitations in previous methods by incorporating user/item semantics and user action importance. Combining two types of noise-reducing filters and multiple training objectives, the HFNF [14] model improves sequential recommendation by capturing user behavior patterns from noisy historical data. AutoSeqRec [32], an efficient sequential recommender system based on autoencoders, captures both long-term user preferences and short-term interests while overcoming limitations in time complexity faced by graph-based methods.

4.4. Attention-based models

To address the limitations of SRSs built on basic neural network structures, some advanced models are combined with certain kinds of basic deep neural networks (e.g., RNN, CNN) to build more powerful SRSs that can address particular challenges. Attention models are commonly employed in SRSs to emphasize those relevant and essential interactions in a sequence while downplaying those irrelevant to the next interaction. BERTSeqHybrid [16] tackles the challenge of personalized POI recommendations for tourists by using BERT and broader user data to overcome the limitations of past sequential models. PAUDRec [11], a new framework for sequential recommendation, tackles popularity bias by considering user desire alongside sequential patterns through a Transformer-based model and separate modules for user desire and popularity effects, achieving better recommendations. Current deep learning recommender systems miss a user's long-term preferences, and FSASA [33] tackles this by combining global and local preference learning modules. MSMT-LSI [34] is a recommendation model that combines multiheaded self-attention and multitemporal embeddings to capture long-term and short-term user preferences and improve recommendation accuracy. DTCoSR [35] addresses existing attention-based sequential recommendation limitations by incorporating collaborative signals and separate self-attention networks for long-term and short-term user interests, achieving better recommendation performance. As the study demonstrates in [36], a new model for recommending the next Point-of-Interest (POI) in location-based social networks by capturing both short-term (context-aware) and long-term user preferences using self-attention and context-aware layers, leading to improved recommendation accuracy and interpretability. MHPE-a [37] combines a temporal point process with an attention mechanism to accurately predict future user interactions by considering both short-term and long-term user preferences from their historical records, ultimately leading to better recommendation performance. SeKeBERT4Rec [38] recommendation model jointly analyzes user session data and keywords to provide more personalized and dynamic recommendations. As a study in [39] shows, they proposed hierarchical attention networks and adversarial learning to address data sparsity and complex group dynamics in group recommender systems, improving individual and group recommendations. In addition, research in [40] introduces BERT4Rec, a new recommendation system model that considers user behavior in both directions (past and future) to capture more nuanced preferences and outperform existing methods. A novel two-layer hierarchical attention network that dynamically learns user long-term and short-term preferences was proposed in [41] from their activity data for sequential recommendations, outperforming existing methods that rely on static representations and linear interactions. The study in [42] investigates the impact of temporal information in sequential recommender systems by identifying and integrating absolute time patterns and relative time patterns into a unified neural model, demonstrating improved performance over existing methods through extensive experiments on real-world datasets.

4.5. Hybrid approaches

These systems combine the strengths of different models, each specializing in capturing distinct types of relationships in user data. The LSRec [8] framework, with its LSIN model, significantly improves recommendation accuracy by utilizing various techniques, such as user–item interaction graphs and user-interest modeling. The SGCross [15] model improves cross-domain recommendation accuracy by capturing user preferences from personal, temporal, and collaborative viewpoints through a multi-view transfer process. OCBSR [17], a new recommender system method, utilizes dynamic user preferences and product image, text, and ID data to significantly outperform existing methods by learning from live streams and short videos. Existing sequential recommenders struggle to consider the meaning/relationships between items and users' long-term preferences. KPHAN [18], a novel method, tackles this limitation by incorporating item relationships from knowledge graphs alongside user long-term and short-term preferences through a hierarchical attention network, achieving significant performance improvements on real-world datasets.

4.6. Contrastive learning-based sequential recommendation

Contrastive loss [23], while gaining traction in sequential recommendation, suffers from limitations such as focusing only on relative distances between positive and negative samples, neglecting absolute constraints and uniform embedding distribution, leading to overly compact embedding spaces, and hindering personalized recommendations. The study in [23] proposes a novel Personalized Contrastive Loss to address these limitations. As the study in [25] shows, leveraging contrastive learning and self-supervised signals in user interaction sequences improves the accuracy of sequential recommendation models beyond the limitations of relying solely on pairwise ranking loss.

4.7. Graph-based models

Recently, with the fast development of GNN, GNN-based SRSs have been devised to leverage GNN to model and capture the complex transitions over user–item interactions in a sequence. Typically, a directed graph is first built on the sequence data by taking each interaction as a node in the graph while each sequence is mapped to a path. Then, the embeddings of users or items are learned on the graph to embed more complex relations over the whole graph. Such an approach fully utilizes GNN's advantage of capturing complex relations in structured relation datasets. GNN-based SRSs have shown great potential to provide explainable recommendations by revealing the complex relations between the recommended items and the corresponding sequential context. However, such SRSs are still in their early stages. As the study demonstrates in [21], the limitations of existing sequential recommendation methods are addressed by proposing a novel framework that leverages dynamic user–item heterogeneous graphs, conditional random fields (CRFs), and the pseudo-likelihood approach to estimate conditional probabilities, ultimately leading to more accurate recommendations that account for both user behavior and user–user influence within a dynamic context. As the study in [22] shows, they proposed a novel graph-based model (TiDA-GCN) that addresses limitations in existing methods for recommending items to users who share accounts across different platforms by considering user interactions across domains, utilizing time intervals between items, and leveraging the underlying graph structure. REDREC [43], a novel recommender system approach, utilizes graph convolutional networks to model item relationships and incorporates time-aware functions to capture the dynamic nature of these relationships for improved sequential recommendations. The research on [44] demonstrates that graph neural network models for both pre-built bundle recommendation and personalized bundle generation achieve significant improvements over existing methods. The study in [45] shows that GCRec considers the order of user interactions and combines short-term and long-term user interests to predict the next item a user will interact with. The research in Hierarchical User Intent Graph Network [46] proposes a new recommender system framework that considers user intent alongside item features to improve the recommendation accuracy for multimedia content like videos and music. The research in [3] develops a Graph Convolution Networks (GCN)-based method to predict user purchase intentions in recommender systems by modeling user price sensitivity and the influence of product categories on price perception, demonstrating its effectiveness through extensive experiments on real-world datasets. Furthermore, the study shown in [47] introduces JMP-GCF, a new recommender system model that considers user preferences for item connectivity and sensitivity to different levels of item popularity for more accurate recommendations. Social conformity bias can skew user data in recommender systems. The TIDE [48] study tackles this by separating the positive influence of genuine item quality and conformity bias, ultimately leading to more accurate recommendations.

5. Evaluation in SRS

This section discusses common evaluation methodologies used in sequential recommendation systems research. Though online evaluation is ideal [49], the proprietary nature of recommender systems and challenges with A/B testing validation necessitate well-designed offline evaluation methods. Current datasets for evaluating sequential recommendation algorithms contain unreliable timestamp information [50], leading to pseudo-sequential data. New datasets with accurate ordering are needed to assess these algorithms better (see Table 2).

5.1. Evaluation datasets

We used four benchmark datasets from three real-world applications for sequential recommendation to evaluate the performance of different models. The datasets vary significantly in domains, platforms, and sparsity:

Amazon[1]: A series of datasets introduced in [52], comprising large corpora of product reviews crawled from Amazon.com. Top-level product categories on Amazon are treated as separate datasets. We consider two categories, 'Beaut' and 'Games.' This dataset is notable for its high sparsity and variability. **Steam:** We introduce a new dataset crawled from Steam, a prominent online video game distribution platform. The dataset contains 2,567,538 users, 15,474 games, and 7,793,069 English reviews.

MovieLens[2]: is a widely used benchmark dataset for evaluating collaborative filtering algorithms. We use the version (MovieLens-1M) that includes 1 million user ratings. We followed the same preprocessing procedure from [52]. For all datasets, we treat a review or rating as implicit feedback (i.e., the user interacted with the item) and use timestamps to determine the sequence order of actions. We discard users and items with less than 5 related actions. For partitioning, we split the historical sequence $S^u$for each user u into three parts: (1) the most recent action $S^u_{|S^u|}$for testing, (2) the second most recent action $S^u_{|S^u|} - 1$ for validation, and (3) all remaining actions for training. Note that the input sequences contain training and validation actions during testing.

5.2. Evaluation metrics

Most of the research on sequential recommendation systems uses the standard ranking accuracy metrics NDCG, MRR, and HitRate.

NDCG (Normalized Discounted Cumulative Gain): NDCG is a measure of ranking quality that considers both the relevance and the position of each result. It is normalized to a value between 0 and 1, where higher values indicate better ranking quality.

MRR (Mean Reciprocal Rank): MRR is a metric that evaluates the effectiveness of rank-ordered retrieval results. It calculates the average of the reciprocal ranks of the first relevant item in the list of retrieved results.

HitRate: HitRate measures the proportion of queries for which the relevant items are present in the top-k ranked results. NDCG at cut-off k:

$$\text{DCG@k} = \sum_{i=1}^{k} \frac{2^{rel_i} - 1}{\log_2(i+1)}$$

$$\text{IDCG@k} = \sum_{i=1}^{|REL|} \frac{2^{rel_i} - 1}{\log_2(i+1)}$$

$$\text{NDCG@k} = \frac{\text{DCG@k}}{\text{IDCG@k}}$$

MRR at cut-off k:

$$\text{MRR@k} = \frac{1}{|Q|} \sum_{q=1}^{|Q|} \frac{1}{\text{rank}_q}$$

HitRate at cut-off k:

$$\text{HitRate@k} = \frac{1}{|Q|} \sum_{q=1}^{|Q|} \text{Hit}_q$$

where

$$\text{Hit}_q = \begin{cases} 1, & \text{if there is at least one relevant item in the top-k results for query } q \\ 0, & \text{otherwise} \end{cases}$$

[1] http://jmcauley.ucsd.edu/data/amazon/

[2] https://grouplens.org/datasets/movielens/

Table 2
Comparison of different modeling methods for sequential recommendation.

Metrics		Model	Year	Modeling method	Data set	Challenges solved	References
HR@10	NDCG@10						
0.73	0.44	RP-SANRec	2024	GRU+Bi-LSTM	Movielens-100k	Dynamic changes in user interest in items over time	[7]
0.74	0.49				Video Game		
0.62	0.38				Beauty		
0.53	0.34	LSRec	2024	LSTM+Transformer encoders	Beauty	Dynamic evolution of user interests over time. Attention mechanisms cannot represent the temporal dimension or use sequence order effectively.	[8]
0.77	0.57				Video Game		
0.88	0.64				Steam		
0.84	0.62				Movielens1M		
0.09	0.04	TFCSRec	2024	Contrastive learning+RNN	Beauty	Existing methods using deep neural networks in the time domain struggle with noisy interactions and sparse data.	[9]
0.03	0.01				Clothing		
0.10	0.05				Toys		
0.08	0.04				Yelp		
0.41	0.28				Gowalla		
	0.05	HFNF	2024	Filter network+adaptive Fourier	Beauty	Noisy interactions in user behavior data can negatively impact the learned representations used for recommendations.	[14]
	0.03				Sports		
	0.01				Clothing		
	0.05				Toys		
	0.05				Steam		
	0.02				ML-1M		
	0.24				Yelp		
	0.31				Gowalla		
0.95	0.49	SGCross	2023	MHG+Attention	CDs-Movies	Methods often focus on a single type of user feature or behavior, neglecting the potential benefits of looking at multiple aspects.	[15]
0.85	0.42				CDs-Movies		
0.76	0.46				Books-Electronics		
0.37	0.21				Electronics-Books		
0.71	0.28				Book-Movie		
0.71	0.17				Movie-Music		
0.69	0.47	OCBSR	2024	Convolutional Autoencoder+Transformer+GRU+ Contrastive Learning	H and M	User preferences for product style, brand, color, etc., can change over time based on exposure to online content.	[16]
0.19	0.08				KuaiRec	Online content offers visual (images) and textual (descriptions) product information, which current systems do not fully utilize.	
	0.01	PAUDRec	2023	Transformer+Causal Graph	Home	Popularity Bias	[17]
	0.02				Sports		
	0.05				Beauty		
	0.07				Office		
0.86	0.59	KPHAN	2024	Knowledge Graph+Attention Network	Movielens-20M	Existing methods only consider user interactions (transitions) and neglect both the meaning/relationships between items and users' long-term preferences.	[11]
0.80	0.55				Movielens-1M		
0.40	0.22				LastFM		

(continued on next page)

Table 2 (*continued*).

Metrics		Model	Year	Modeling method	Data set	Challenges solved	References
HR@10	NDCG@10						
0.05	0.03	MGT	2023	Transformer	Beauty	Existing sequential recommenders struggle to capture the collective influence of multiple past items, only considering individual item effects, and MGT tackles this limitation.	[18]
0.01	0.008				Clothing		
0.32	0.18				ML-1M		
	0.52	T3-BERT4Rec	2024	Transformer	Steam	Transformer-based recommenders suffer from overfitting.	[12]
	0.37				ML-1M		
	0.30				ML-20M		
18.45	9.96	GaGNN	2024	GNN	MovieLens-100K	Current sequential recommenders miss group behavior and user data limitations. GaGNN tackles this by incorporating collaborative graphs and multi-source information fusion.	[13]
23.49	12.45				MovieLens-1M		
6.68	3.89				Beauty		
4.3	4.75				Cell Phone		
9.11	5.71				LibraryThing		
9.68	5.88				CD		
0.45	0.26	TGT	2022	Temporal Graph+ Transformer	Taobao Data	Existing sequential recommenders miss the richness of multi-typed user interactions and their dependencies.	[51]
0.51	0.33				IJCAI Contest		
0.09	0.05	PCL4SRec	2023	CL+BCE/CE	Beauty	Existing sequential recommendation methods neglect the importance of the loss function. CL focuses solely on pushing positive and negative samples apart without absolute constraints or ensuring uniform distribution of all item embeddings.	[19]
0.05	0.03				Sports		
0.09	0.05				Toys		
0.26	0.14				ML-1M		
0.46		HGESNN	2022	GNN	Industrial	Existing CTR prediction models for recommender systems tend to focus on highly relevant items based on user history. Leading to limited exploration and under-exposure of potentially interesting items.	[23]
0.05	4.8				ML-20M		
0.48	0.32	SASRec	2018	Attention mechanism	Beauty	The trade-off between capturing long-term user behavior and computational efficiency.	[25]
0.74	0.53				Games		
0.87	0.63				Steam		
0.82	0.59				MovieLens-1M		
0.24	0.14	ContraRec	2023	CL	Beauty	Existing sequential recommendation methods primarily rely on pairwise ranking loss. Neglecting the potential of contrastive learning and self-supervised signals within user interaction sequences.	[24]
0.42	0.25				Yelp-2018		
0.78	0.59				Gowalla		

6. Results and discussion

6.1. Compared model

We compared the models that have performed well in the history of sequential recommendation systems—the compared model on the same data set and hyperparameters.

GRU4Rec[3] [53]: GRU4Rec is an RNN-based session recommendation method that adopts GRU to model user sequences and achieves good results. **GRU4Rec+**[4] [54]: GRU4Rec+ is an improved version of GRU4Rec that addresses some of the shortcomings of the original version and achieves significant performance improvements. **Caser**[5] [31]: Caser embeds a sequence of recent items into an "image" in time and latent space and uses convolutional filters to learn sequential patterns from the "image".

SASRec[6] [52]: SASRec is a state-of-the-art sequential recommendation method that uses the self-attention mechanism.

[3] https://github.com/hidasib/GRU4Rec

[4] https://github.com/hidasib/GRU4Rec

[5] https://github.com/graytowne/caser

[6] https://github.com/kang205/SASRec

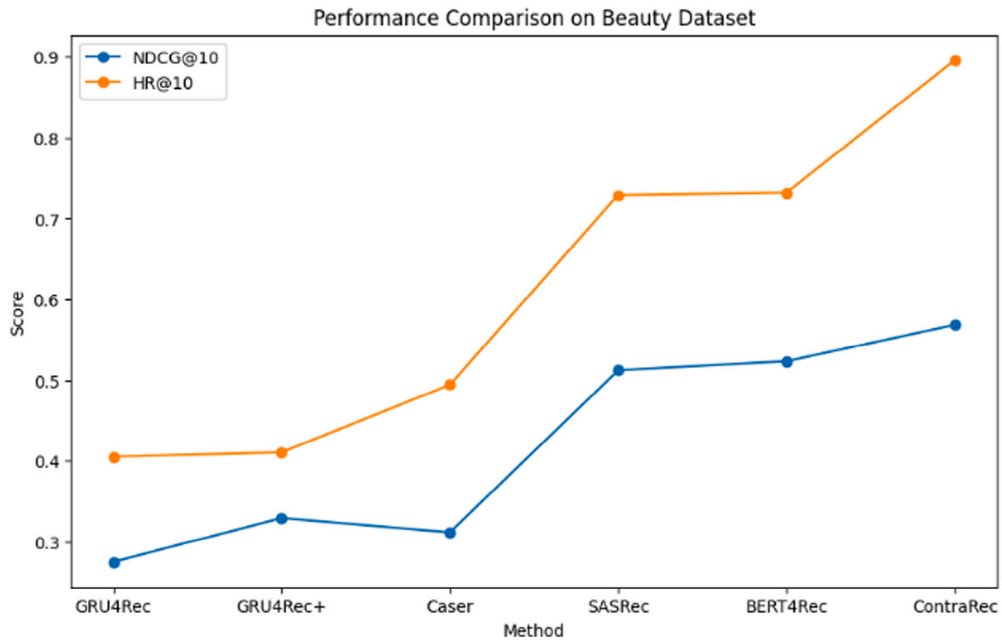

Fig. 2. Model performance comparison.

Table 3
Performance comparison on Beauty dataset.

Method	Beauty	
	NDCG@10	HR@10
GRU4Rec	0.2752	0.4054
GRU4Rec+	0.3296	0.4108
Caser	0.3119	0.4944
SASRec	0.5124	0.7293
BERT4Rec	0.5235	0.7324
ContraRec	0.5687	0.8962

ContraRec [25]: a framework for sequential recommendation that unifies context-target and context–context contrastive learning tasks, demonstrating superior performance compared to state-of-the-art methods by leveraging self-supervised signals and integrating various sequential recommendation models.

BERT4Rec [40]: a sequential recommendation model that addresses limitations of unidirectional architectures by using deep bidirectional self-attention and the Cloze objective to model user behavior sequences, demonstrating superior performance over state-of-the-art methods in extensive experiments.

The line graph on Fig. 2 and Table 3 illustrates the performance comparison of various sequential recommendation methods on the Beauty dataset. The metrics used for evaluation are NDCG@10 (Normalized Discounted Cumulative Gain at rank 10) and HR@10 (Hit Rate at rank 10). The graph shows that ContraRec achieves the highest performance in both NDCG@10 and HR@10, indicating its superior ability to capture user preferences and recommend relevant items. BERT4Rec also performs notably well, closely following ContraRec, demonstrating the effectiveness of using bidirectional self-attention mechanisms to model user behavior sequences. SASRec, another strong performer, leverages self-attention mechanisms, achieving high scores in both metrics but slightly lower than BERT4Rec and ContraRec. Caser, which uses convolutional filters to learn sequential patterns, shows moderate performance improvement over the earlier GRU4Rec and its improved version, GRU4Rec+. Although GRU4Rec+ outperforms GRU4Rec, both methods lag behind the more advanced techniques like SASRec, BERT4Rec, and ContraRec. The choice of these systems for comparison is justified by their diverse approaches to modeling sequential data. GRU4Rec and GRU4Rec+ are based on recurrent neural networks and are among the first models applied to session-based recommendations. Caser introduces convolutional neural networks to capture patterns in sequences. SASRec, BERT4Rec, and ContraRec represent the latest advancements, utilizing self-attention mechanisms and contrastive learning for more accurate recommendations. The results generally indicate methods incorporating self-attention and contrastive learning, such as ContraRec and BERT4Rec, significantly outperform traditional RNN-based methods like GRU4Rec and its variants. This highlights the importance of adopting advanced modeling techniques to improve recommendation system performance.

6.2. Hardware and software configuration

The experiments were conducted on a server with the following specifications:

The computational environment for the experiments comprises several key specifications. The CPU used is an Intel Xeon E5-2690 v4 @ 2.60 GHz, providing robust processing capabilities. The system is equipped with 256 GB of RAM, ensuring ample memory for handling large datasets and complex computations. For accelerated computing, the environment includes 2 *x* NVIDIA Tesla V100 GPUs, which are highly suitable for deep learning tasks. The Operating System is Ubuntu 20.04 LTS, offering a stable and secure platform for running various applications. The main Software tools utilized are Python 3.8, PyTorch 1.8, and CUDA 11.2, forming a powerful stack for machine learning and deep learning development

6.3. Model training

The models were trained using the Adam optimizer with an initial learning rate of 0.001. The batch size was set to 128, and training was conducted for 50 epochs. Early stopping with a patience of 5 epochs was employed to prevent overfitting.

6.4. Hyperparameter tuning

Hyperparameters for all models were tuned using grid search over the following ranges: The hyperparameters used in the experiments include a range of values for several key settings. The Learning Rate is tested at three different levels: 0.0001, 0.001, and 0.01, to determine the optimal rate for model convergence. The Batch Size is varied among 64, 128, and 256, allowing for experimentation with different levels of gradient update frequencies. Additionally, the Dropout Rate is set at 0.1, 0.2, and 0.3 to assess the impact of regularization on model performance. The best hyperparameters were selected based on the validation set performance.

7. Open issues and future directions

In recent years, particularly the last three years, we have witnessed the fast development of sequential recommender systems and the prosperity of deep learning, especially recurrent neural networks. While categorizing and summarizing the research practices in this field, we have identified further open research directions discussed below.

7.1. Context-aware sequential recommender systems

While traditional sequential recommender systems focus solely on the sequence of user interactions, context-aware sequential recommender systems represent a significant advancement. They delve deeper, considering the sequence of what a user interacts with and the broader context in which those interactions occur. This context encompasses a diverse range of factors, such as:

Time: When did the interaction happen? Was it on a weekday, weekend, or holiday? Time of day can also be relevant (e.g., morning browsing might suggest interest in news or breakfast items, while evening browsing might suggest interest in entertainment or dinner options).

Location: Where were the users when they interacted with the item? It could be their physical location (e.g., home, work, gym) or a virtual location within an app (e.g., browsing a specific category).

User Device: What device was the user using to interact (e.g., phone, laptop, smart speaker)? It can influence the types of items they might be interested in (e.g., a phone user might be more likely to stream music, while a laptop user might be more likely to research a product).

User Attributes: User Attributes are not just demographics like age, gender, or interests. They also include more dynamic factors like the user's current mood or recent purchases, which can significantly influence the recommendations provided.

7.2. Social-aware sequential recommender systems

Many recommender systems currently focus solely on a user's individual interaction history. However, a growing area of research is exploring social-aware sequential recommender systems. These systems recognize that users are influenced by the social circles they interact with, both online and offline. Social-aware sequential recommender systems represent a promising future direction. They can provide more relevant recommendations by leveraging the power of social influence and social connections.

7.3. Group sequential recommendation

Group recommendation systems aim to suggest items for a group of users, considering all members' preferences. However, a significant challenge in this area is data sparsity. This means there is often limited data on how groups interact with items.

7.4. Model both long-term user/group preferences and short-term trends

Group recommender systems aim to suggest items that cater to a group's collective interests. However, user and group preferences and item trends can be dynamic (changing frequently) or static (relatively stable over time). Sequential recommender systems must consider both aspects to make truly effective group recommendation sets.

7.5. Cross-domain recommendation with time-sequential data

Recommender systems traditionally focus on a user's activity within a single domain, like movies or music. However, exploring cross-domain recommendation tasks with time-sequential data is a promising future direction. This approach aims to leverage user activity across domains to make more informed recommendations.

8. Conclusion

In this survey paper, we provided an overview of sequential recommendation systems, including methodologies, evaluation techniques, and future research directions. We discussed various approaches for modeling sequential user behavior and highlighted the importance of considering temporal dynamics in recommendation systems. Despite significant advancements in the field, there remain challenges and opportunities for further research to enhance the effectiveness and scalability of sequential recommendation systems. Furthermore, we have provided a comprehensive examination of sequential recommendation systems (SRS), emphasizing their growing importance in effectively capturing and predicting users' evolving preferences over time. Unlike traditional recommender systems that primarily focus on static user–item interactions, SRS models take into account the sequential order of interactions, enabling more dynamic and personalized recommendations. We categorized and discussed various approaches, including traditional methods like collaborative filtering and content-based filtering, and advanced techniques such as attention based model, and deep learning models. Our exploration highlights the strengths and limitations of each approach, with deep learning techniques, particularly Recurrent Neural Networks (RNNs) and their variants like LSTMs and GRUs, showing significant promise in handling complex sequential data. Additionally, attention-based models and hybrid approaches have emerged as powerful tools for capturing nuanced user behavior and preferences, demonstrating superior performance in various application domains. We also discussed the key concepts essential for understanding SRS, such as user–item interactions, sequence modeling, and the distinction between session-based and long-term preferences. Evaluation methodologies for these systems were reviewed, underscoring the importance of robust metrics to assess their effectiveness accurately. Despite the advancements in SRS, several challenges and opportunities remain. Issues such as data sparsity, scalability, and the need for more sophisticated models to capture long-term dependencies continue to drive research in this field. Future directions include developing more efficient algorithms, leveraging contrastive learning and graph-based models, and exploring new hybrid approaches that integrate various techniques to enhance recommendation accuracy further. In conclusion, sequential recommender systems represent a significant leap forward in personalized recommendation technology. By continuously evolving to better understand and predict user behavior, these systems hold great potential to improve user experiences across diverse applications. Continued research and innovation in this area will undoubtedly lead to even more sophisticated and effective recommendation systems in the future.

CRediT authorship contribution statement

Tesfaye Fenta Boka: Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Zhendong Niu:** Supervision, Resources. **Rama Bastola Neupane:** Writing – review & editing, Data curation.

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.

Data availability

Data will be made available on request.

Acknowledgments

We gratefully acknowledge the support provided by the Beijing Institute of Technology Computer Science Department, China for this work. Special thanks are extended to Professor Zhendong Niu from the School of Computer Science and Technology, China, for his invaluable guidance and meticulous proofreading of this paper. His expertise and insights have significantly enhanced the quality of our manuscript. We appreciate the time and effort devoted to ensuring the clarity and precision of our work

References

[1] L. Wu, X. He, X. Wang, K. Zhang, M. Wang, A survey on accuracy-oriented neural recommendation: From collaborative filtering to information-rich recommendation, IEEE Trans. Knowl. Data Eng. 35 (5) (2022) 4425–4445.

[2] W. Liao, Q. Zhang, B. Yuan, G. Zhang, J. Lu, Heterogeneous multidomain recommender system through adversarial learning, IEEE Trans. Neural Netw. Learn. Syst. (2022).

[3] Y. Zheng, C. Gao, X. He, D. Jin, Y. Li, Incorporating price into recommendation with graph convolutional networks, IEEE Trans. Knowl. Data Eng. 35 (2) (2021) 1609–1623.

[4] M. Quadrana, P. Cremonesi, D. Jannach, Sequence-aware recommender systems, ACM Comput. Surv. (CSUR) 51 (4) (2018) 1–36.

[5] B. Wang, S. Song, S. Liu, X. Deng, MultiFDF: Multi-community clustering for fairness-aware recommendation, IEEE Trans. Comput. Soc. Syst. (2023).

[6] D. Jannach, P. Pu, F. Ricci, M. Zanker, Recommender systems: Past, present, future, Ai Mag. 42 (3) (2021) 3–6.

[7] X. Zheng, X. Li, Z. Chen, L. Sun, Q. Yu, L. Guo, Y. Luo, Enhanced self-attention mechanism for long and short term sequential recommendation models, IEEE Trans. Emerg. Top. Comput. Intell. (2024).

[8] W. Mu, Q. Liu, H. Cheng, M. Zhuo, Long-short interest network with graph-based method for sequential recommendation, Int. J. Mach. Learn. Cybern. (2024) 1–13.

[9] Y. Xiao, J. Huang, J. Yang, TFCSRec: Time–frequency consistency based contrastive learning for sequential recommendation, Expert Syst. Appl. 245 (2024) 123118.

[10] B. Wu, X. Su, J. Liang, Z. Sun, L. Zhong, Y. Ye, Graph gating-mixer for sequential recommendation, Expert Syst. Appl. 238 (2024) 122060.

[11] J. Wu, B. Yang, R. Mao, Q. Li, Popularity-aware sequential recommendation with user desire, Expert Syst. Appl. 237 (2024) 121429.

[12] Y. Zhang, B. Yang, R. Mao, Q. Li, MGT: Multi-granularity transformer leveraging multi-level relation for sequential recommendation, Expert Syst. Appl. 238 (2024) 121808.

[13] H. Li, J. Zhao, H. Huo, S. Fang, J. Chen, L. Yao, Y. Hua, T3SRS: Tensor train transformer for compressing sequential recommender systems, Expert Syst. Appl. 238 (2024) 122260.

[14] Y. Xiao, J. Huang, J. Yang, HFNF: learning a hybrid Fourier neural filter with a heterogeneous loss for sequential recommendation, Appl. Intell. 54 (1) (2024) 283–300.

[15] H. Li, L. Yu, X. Niu, Y. Leng, Q. Du, Sequential and graphical cross-domain recommendations with a multi-view hierarchical transfer gate, ACM Trans. Knowl. Discov. Data 18 (1) (2023) 1–28.

[16] A. Noorian, A BERT-based sequential POI recommender system in social media, Comput. Stand. Interfaces 87 (2024) 103766.

[17] Y. Lu, Y. Duan, Online content-based sequential recommendation considering multimodal contrastive representation and dynamic preferences, Neural Comput. Appl. (2024) 1–19.

[18] S. Ruan, C. Yang, D. Li, Knowledge-enhanced personalized hierarchical attention network for sequential recommendation, World Wide Web 27 (1) (2024) 2.

[19] L. Xia, C. Huang, Y. Xu, J. Pei, Multi-behavior sequential recommendation with temporal graph transformer, IEEE Trans. Knowl. Data Eng. (2022).

[20] T. Zhu, L. Sun, G. Chen, Graph-based embedding smoothing for sequential recommendation, IEEE Trans. Knowl. Data Eng. 35 (1) (2021) 496–508.

[21] Z. Li, R. Cai, F. Wu, S. Zhang, H. Gu, Y. Hao, Y. Yan, TEA: A sequential recommendation framework via temporally evolving aggregations, IEEE Trans. Neural Netw. Learn. Syst. (2022).

[22] L. Guo, J. Zhang, L. Tang, T. Chen, L. Zhu, H. Yin, Time interval-enhanced graph neural network for shared-account cross-domain sequential recommendation, IEEE Trans. Neural Netw. Learn. Syst. (2022).

[23] H. Tang, G. Zhao, J. Gao, X. Qian, Personalized representation with contrastive loss for recommendation systems, IEEE Trans. Multimed. (2023).

[24] J. Chen, Z. Gong, Y. Li, H. Zhang, H. Yu, J. Zhu, G. Fan, X.-M. Wu, K. Wu, Meta-path based neighbors for behavioral target generalization in sequential recommendation, IEEE Trans. Netw. Sci. Eng. 9 (3) (2022) 1658–1667.

[25] C. Wang, W. Ma, C. Chen, M. Zhang, Y. Liu, S. Ma, Sequential recommendation with multiple contrast signals, ACM Trans. Inf. Syst. 41 (1) (2023) 1–27.

[26] S. Wang, L. Hu, Y. Wang, L. Cao, Q.Z. Sheng, M. Orgun, Sequential recommender systems: challenges, progress and prospects, 2019, arXiv preprint arXiv:2001.04830.

[27] S. Rendle, W. Krichene, L. Zhang, Y. Koren, Revisiting the performance of ials on item recommendation benchmarks, in: Proceedings of the 16th ACM Conference on Recommender Systems, 2022, pp. 427–435.

[28] S. Rendle, C. Freudenthaler, L. Schmidt-Thieme, Factorizing personalized markov chains for next-basket recommendation, in: Proceedings of the 19th International Conference on World Wide Web, 2010, pp. 811–820.

[29] R. He, J. McAuley, Fusing similarity models with markov chains for sparse sequential recommendation, in: 2016 IEEE 16th International Conference on Data Mining, ICDM, IEEE, 2016, pp. 191–200.

[30] D. Guo, C. Wang, Sequence recommendation based on deep learning, Comput. Intell. 36 (4) (2020) 1704–1722.

[31] J. Tang, K. Wang, Personalized top-n sequential recommendation via convolutional sequence embedding, in: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining, 2018, pp. 565–573.

[32] S. Liu, J. Liu, H. Gu, D. Li, T. Lu, P. Zhang, N. Gu, Autoseqrec: Autoencoder for efficient sequential recommendation, in: Proceedings of the 32nd ACM International Conference on Information and Knowledge Management, 2023, pp. 1493–1502.

[33] S. Guo, X. Liao, F. Meng, Q. Zhao, Y. Tang, H. Li, Q. Zong, FSASA: Sequential recommendation based on fusing session-aware models and self-attention networks, Comput. Science. Inf. Syst. (2023) 67.

[34] H. Wang, Y. Wang, Y. Liu, A sequential recommendation model for balancing long-and short-term benefits, Int. J. Comput. Intell. Syst. 17 (1) (2024) 75.

[35] L. Liu, Dynamic time-aware collaborative sequential recommendation with attention-based network, Knowl. Inf. Syst. 66 (3) (2024) 1639–1655.

[36] T. Thaipisutikul, Y.-N. Chen, An improved deep sequential model for context-aware POI recommendation, Multimedia Tools Appl. 83 (1) (2024) 1643–1668.

[37] D. Wang, X. Zhang, Z. Xiang, D. Yu, G. Xu, S. Deng, Sequential recommendation based on multivariate hawkes process embedding with attention, IEEE Trans. Cybern. 52 (11) (2021) 11893–11905.

[38] H. Veeramani, S. Thapa, U. Naseem, Temporally dynamic session-keyword aware sequential recommendation system, in: 2023 IEEE International Conference on Data Mining Workshops, ICDMW, IEEE, 2023, pp. 157–164.

[39] R. Liang, Q. Zhang, J. Wang, J. Lu, A hierarchical attention network for cross-domain group recommendation, IEEE Trans. Neural Netw. Learn. Syst. (2022).

[40] F. Sun, J. Liu, J. Wu, C. Pei, X. Lin, W. Ou, P. Jiang, BERT4rec: Sequential recommendation with bidirectional encoder representations from transformer, in: Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 2019, pp. 1441–1450.

[41] H. Ying, F. Zhuang, F. Zhang, Y. Liu, G. Xu, X. Xie, H. Xiong, J. Wu, Sequential recommender system based on hierarchical attention network, in: IJCAI International Joint Conference on Artificial Intelligence, 2018.

[42] W. Ye, S. Wang, X. Chen, X. Wang, Z. Qin, D. Yin, Time matters: Sequential recommendation with complex temporal information, in: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2020, pp. 1459–1468.

[43] R. Yao, W. Xu, Z. Liu, Y. Wang, Z. Li, Y. Qiao, J. Yang, Redrec: Relation and dynamic aware graph convolutional network for sequential recommendation, in: 2023 8th IEEE International Conference on Network Intelligence and Digital Content, IC-NIDC, IEEE, 2023, pp. 192–196.

[44] J. Chang, C. Gao, X. He, D. Jin, Y. Li, Bundle recommendation and generation with graph neural networks, IEEE Trans. Knowl. Data Eng. 35 (3) (2021) 2326–2340.

[45] B. Wu, X. He, Q. Zhang, M. Wang, Y. Ye, GCRec: Graph-augmented capsule network for next-item recommendation, IEEE Trans. Neural Netw. Learn. Syst. (2022).

[46] Y. Wei, X. Wang, X. He, L. Nie, Y. Rui, T.-S. Chua, Hierarchical user intent graph network for multimedia recommendation, IEEE Trans. Multimed. 24 (2021) 2701–2712.

[47] K. Liu, F. Xue, X. He, D. Guo, R. Hong, Joint multi-grained popularity-aware graph convolution collaborative filtering for recommendation, IEEE Trans. Comput. Soc. Syst. 10 (1) (2022) 72–83.

[48] Z. Zhao, J. Chen, S. Zhou, X. He, X. Cao, F. Zhang, W. Wu, Popularity bias is not always evil: Disentangling benign and harmful bias for recommendation, IEEE Trans. Knowl. Data Eng. (2022).

[49] B. Hidasi, Á.T. Czapp, Widespread flaws in offline evaluation of recommender systems, in: Proceedings of the 17th ACM Conference on Recommender Systems, 2023, pp. 848–855.

[50] D. Woolridge, S. Wilner, M. Glick, Sequence or pseudo-sequence?, 2021.
[51] Z. Huang, Z. Sun, J. Liu, Y. Ye, Group-aware graph neural networks for sequential recommendation, Inform. Sci. (2024) 120623.
[52] W.-C. Kang, J. McAuley, Self-attentive sequential recommendation, in: 2018 IEEE International Conference on Data Mining, ICDM, IEEE, 2018, pp. 197–206.
[53] B. Hidasi, A. Karatzoglou, L. Baltrunas, D. Tikk, Session-based recommendations with recurrent neural networks, 2015, arXiv preprint arXiv:1511.06939.
[54] B. Hidasi, A. Karatzoglou, Recurrent neural networks with top-k gains for session-based recommendations, in: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, 2018, pp. 843–852.