Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 13, No. 3, 2024

Cold-Start Product Recommendation Method Based on GAE

Bo He [a], Jingxuan Zhou *, Ruoyu Zhao [b], Lei Wang [c], Longbing Li [d], Yongfen Yang [e]

School of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China

* **Corresponding author**: Jingxuan Zhou (Email: zjx@stu.cqut.edu.cn), [a] hebo@cqut.edu.cn, [b] qiaqiaeric@outlook.com, [c]wanglei6731@163.com, [d]llbgod@163.com, [e]yangff0518@163.com

Abstract: Recommendation systems play a pivotal role in alleviating information overload, delivering customized advisory services, and aiding users in making investment decisions. However, the cold start problem in recommendation systems has always been in urgent need of solution and optimization, especially for new item uptake or new user engagement. Owing to this, this paper categorizes the conventional and cutting-edge methodologies for addressing the cold start problem, and elucidates the research progress and outstanding methods in recent years. Firstly, this paper provides an exhaustive review of the various strategies and methods proposed by researchers to alleviate the cold start problem. Secondly, the paper summarizes several traditional recommendation methods for mitigating cold-start recommendations. Thirdly, the paper synthesizes the recent strategies and approaches into two categories: data-driven strategies and approach-driven strategies. Furthermore, the approach-driven strategies are categorized into five main clusters based on meta learning, session, heterogeneous graph and attributed graph, and other novel approaches, Furthermore, the method-driven strategies are refined into five primary clusters: meta-learning, session-based learning, heterogeneous graphs, attributed graphs, and other innovative methods. Particularly, a model named GAE, composed of graph attention neural networks (GAT) and meta-learning, is highlighted. This model initially aggregates and extracts attribute feature information from the neighbor nodes of cold-start users using the GAT; then, it utilizes the Embedding Generator (EG) to consider the cold-start users' own attribute information, combining it with the associated neighbor attribute feature information to generate the final attribute feature information for the cold-start users; meta-learning is employed to generate a list of recommended items for the cold-start users. Finally, the possible research directions to solve the cold start problem in the future are pointed out, which include complex relationship learning, multimodal recommendation, and privacy protection.

Keywords: Recommendation Systems; Cold Start Problem; Meta Learning; Heterogeneous Graph; Attributed Graph.

1. Introduction

The immense increase of online information has led to the phenomenon of "information overload," where users struggle to identify and access valuable content amidst the vast sea of data. Recommendation systems serve as intelligent filtering tools that tailor recommendations to users' preferences and interests, facilitating the discovery and acquisition of relevant information. By doing so, they enhance information retrieval efficiency and user experience. Over the past few decades, recommendation systems have gradually evolved, with recommendation algorithms being continuously updated. These algorithms can be classified into three major types based on the techniques used for filtering items: content-based, collaborative filtering, and hybrid recommendation systems. Similarly based on the type or source of data, the recommendation systems can also be of five types: social network-based recommendation system, knowledge-based recommendation systems, utility-based recommendation systems, context-aware recommendation systems, and demographic-based recommendation systems. Beyond these well-established categories, the evolving research has given rise to several other variants of knowledge-based recommender systems that employ various data sources to derive knowledge for modeling. For instance, critiquing-based recommender systems utilize user feedback as critiques to make item recommendations. Among all recommender systems, collaborative filtering and content-based approaches are the most widely adopted and classic recommendation techniques. The collaborative filtering method leverages past collaborative preferences between the items and users to generate top N item recommendations for users. Several collaborative filtering approaches are renowned for their higher accuracy; however, they encounter cold-start user and cold-start item problems when faced with new users and new items, respectively. These issues affect the prediction accuracy and quality of recommendations. In contrast, content-based recommender systems overcome these cold-start scenarios by analyzing item metadata and identifying similarities. However, they still face challenges in providing reliable recommendations in the absence of user interaction data.

Indeed, the expanding scale of recommendation systems and the broadening scope of recommended items have magnified the prominence of the cold-start problem. This challenge emerges when new users or items enter the system, lacking adequate information for similarity or association matching. Consequently, the recommendation system is unable to generate accurate recommendations, thereby resulting in a cold-start situation. Cold Start problems also pose serious challenges to the business value of recommendation systems. New users who will not find the recommendations useful are likely to discontinue using the system, thereby negatively impacting user engagement. The cold-start issue for new items also escalates operational and marketing costs for businesses, as promoting these new products necessitates substantial time and financial investments in the absence of organic recommendation methods. Hence, to alleviate the challenges posed by the cold-start problem for new users or items, several researchers from both academia and industry have actively engaged in designing, developing, and proposing novel strategies. The aim of this paper is to organize and analyze effective strategies for alleviating the cold-start problem, providing a concise overview of the field.

2. Research Objective and Related Works

The cold start problem presents a significant challenge to the efficacy and performance of recommender systems. Consequently, a multitude of approaches and algorithms have been proffered by researchers across both academic and industrial domains. However, there has been very limited recent research done on collating these approaches and algorithms. To address this gap, the paper conducts a systematic literature review of various strategies and approaches proposed by researchers in the last decade. This research observes few notable works in this area.

First, MAO conducted a comprehensive classification of various methods used for cold-start recommendation, dividing them into data-driven and method-driven approaches, and provided strategies for mitigating cold-start issues in specific domains.

Second, Surana and Basari conducted a literature review aiming to identify the key components of recommendation systems, existing methods, and how to measure their effectiveness. However, their study, which covered papers from 2006 to 2016, did not employ a systematic literature review method. Consequently, the strategies discussed may not have exhaustive coverage of all possible approaches and algorithms, potentially missing out on some of the latest research findings. Additionally, the authors acknowledge that the study did not consider all application areas of recommendation systems, potentially overlooking certain domains, such as online shopping applications.

Lastly, Camacho and Alves-Souza conducted a systematic review examining the utilization of social network data to address the cold start problem in recommender systems. This research, however, specifically targets collaborative filtering-based recommender systems, aiming to explore the effectiveness of integrating social network data within various CF methodologies in mitigating the cold start challenge.

Building upon these research efforts, this paper examines existing studies on the cold-start problem, with the objective of compiling the various methods and algorithms that have been proposed to mitigate these issues in the contexts of new users and new items. Thus, the research objectives of this study are:

(1) To conduct an exhaustive study using systematic literature review to identify the existing methods and algorithms that have been employed to alleviate the cold-start problem.

(2) To synthesizes the various strategies adopted by researchers across different domains of recommender systems, including benchmark algorithms, types of information and their effectiveness.

The remainder of the paper is organized as follows. Section 3 discusses traditional methods for addressing the cold-start problem. in this study. The literature review is further elaborated in Section 4, where a qualitative synthesis is also presented, categorizing the strategies and methods proposed by researchers to address the cold-start problem. Section 5 further discusses the synthesis approach and offers recommendations for future research. Finally, Section 6 summarizes the findings and contributions of this paper.

3. Traditional Approaches for Addressing the Cold Start Problem

According to the existing research on the cold-start problem, traditional recommendation algorithms primarily address this issue from three main angles: content-based, collaborative filtering-based, and hybrid recommendation systems-based approaches.

3.1. Content-Based Recommendation Approaches

The core principle of content-based recommendation (CB) techniques is to generate personalized recommendations by leveraging the alignment between item features and user preferences. Chang applied recommendation systems to the domain of homestay family recommendations, proposing the creation of a Chinese word segmentation corpus utilizing a vast collection of user reviews on homestays. The corpus was trained with word2vec to extract label features of homestay families. Subsequently, TF-IDF was employed to vectorize the label features, and the cosine similarity was calculated to determine the similarity of feature vectors, thereby deriving the recommendation results for new users. To address the issue of neglecting learner characteristics in e-learning recommendation systems, Jeevamol proposed an ontology-based content recommendation system model. In the context of similarity analysis, Ana compared distance metrics and evaluated the similarity of music using four measures: Minkowski distance, Euclidean distance, Manhattan distance, and Bray-Curtis similarity. They aimed to identify feature extraction and engineering methods suitable for the background of classical music similarity analysis, and to determine the performance differences among the selected distance metrics, considering the multidimensional nature of music. Their work contributes to the construction of a recommendation system for classical music. However, content-based recommendation algorithms primarily address the cold-start problem for new items and are less effective in resolving the cold-start issue for new users.

3.2. Collaborative Filtering-Based Recommendation Approaches

Collaborative filtering (CF) algorithms are among the most successful and widely adopted methods in recommender systems. The core principle of CF algorithms is to analyze the interaction behaviors of different users with items, inferring that users who exhibit similar interactions with items are likely to have similar preferences. Essentially, CF algorithms tackle a matrix completion problem. Fan propose an enhanced collaborative filtering method that integrates user preferences and trust degrees for recommendation. Wu utilized users' attribute information and employed a K-means clustering algorithm optimized based on user attribute features to perform clustering, generating multiple clusters. By integrating the user attribute features within each cluster, they established a new similarity computation model. Through searching for nearest neighbors within the clusters, they were able to generate recommendation lists to realize recommendations. Xue modeled the nonlinear and high-order relationships between items using a nonlinear neural network, thereby uncovering more underlying factors in user decision-making.

3.3. Hybrid Recommendation Systems Based Approaches

Hybrid recommendation (HR) methods involve the combination of multiple recommendation techniques to achieve higher recommendation accuracy and to mitigate issues such as cold start and data sparsity that may arise with individual recommendation technologies. Esteban proposed a course recommendation method that integrates multiple recommendation strategies using a genetic algorithm. This method applies a customized genetic algorithm in the pre-recommendation stage to optimize the parameter configuration of the recommendation system using training data, and then uses this configuration to construct the recommendation system model. In the practice question recommendation method proposed by Wu, a combination of session-based recommendation methods and the simulated annealing algorithm is employed to enhance recommendation diversity and novelty while ensuring the accuracy of the recommendations. Hu introduced a hybrid recommendation algorithm that combines the latent factor model (LFM) with the personal rank (PR) algorithm based on graphs to enhance the recommendation process.

A comparative analysis of traditional recommendation algorithms is conducted, and their respective advantages and disadvantages are summarized (refer to Table 1).

Table 1. Comparison of advantages and disadvantages of traditional recommendation technology

Approaches	Advantages	Disadvantages
CB	1. Interpretability	1. Poor feature extraction performance
	2. Ease for implementation	2. Poor security
CF	1.Easy to operate	1. Unable to handle complex recommendations
	2. Ease of modeling	2. Lack of interpretability
HR	1. Overcomes data sparsity	1. Model complexity, low efficiency
	2. Suits large-scale user recommendations	2. Complex recommendation process
		3. Poor interpretability

4. Recent Algorithms for Addressing the Cold Start Problem

Traditional recommendation methods have become inadequate to meet the diverse recommendation needs of users. Therefore, numerous new methods have been proposed to address this issue.

4.1. Data-driven Strategies

Data-driven strategies proposed by various researchers have employed unique methods to address the cold-start problem. In addition to utilizing user feedback and item information, these strategies also leverage other types of data, such as social networks, trust, location, and cross-domain data to handle the cold start problem, including user demographic data, cross-domain data, and social network data commonly.

4.1.1. Demographic Data Based Approaches

Raigoza utilized demographic characteristics to categorize users into different groups, and then recommended items based on their popularity or relevance within user groups with similar characteristics.

Lika on the basis of classifying users by demographic features, designed a rating prediction function, and made recommendations for new users based on the ratings of users in the same group.

Mu proposed a collaborative recommendation model based on an auxiliary stacked denoising autoencoder (ASDAE), which integrates auxiliary information with rating information. The model learns user preferences from both auxiliary and rating information to optimize the cold-start problem.

Cold-start solutions based on demographic data can only address the cold-start problem for new users and are not capable of resolving the cold-start issue for new items.

4.1.2. Social Network Based Approaches

Social network data of users is typically employed in the context of recommender systems to enhance the rating profiles of users who lack sufficient ratings and trust information. XU proposed a matrix factorization model that integrates users' social networks, as illustrated in Figure 1. The model clusters users using a clustering algorithm based on particle swarm optimization and then employs matrix factorization to calculate user preferences. Although this method effectively mitigates the cold-start problem, the training complexity increases when the training data is large. Ahmadian have proposed a “Reputation-based Trust-Aware Recommender System (RTARS)” that makes use of social information of users to enhance the rating profiles of users that do not have sufficient ratings and trust information.

Researchers have also coupled social network data with other data attributes to effectively remediate new user cold start problems. Ojagh proposed a User Similarity Detection Engine (USDE) that leverages users' personal intelligent devices to extract their social information, enabling the detection of user similarity and the generation of recommendations.

Herce-Zelaya proposed a method that creates behavioral profiles using social media data, classifies users, and then uses classification trees and random forests to make predictions for these users.

Ghaviour and Meybodi leveraged a trust network replete with trust (highly reputable popular users) and interest similarity to introduce a random trust-based propagation method, LTRS, with the objective of improving the quality and coverage of recommendations while also addressing the challenges of cold start and data sparsity.

4.1.3. Cross Domain Based Approaches

Cross-domain recommendation is a method that leverages useful information from source domains to compensate for the deficiencies in target domain content. This method integrates behavioral data of users across multiple domains, combining user preference characteristics from different domains, effectively addressing data sparsity and cold start issues. Moreover, it contributes to enhancing the accuracy and diversity of recommendations.

Wang calculate user similarity based on user rating behavior and embed it into a matrix factorization model as the foundation for cross-domain knowledge transfer. Subsequently, they propose a neighbor-based cross-domain latent feature mapping method. For cold-start users, this method allows for learning user features based on their neighbors.

Wang proposed a cross-domain latent feature mapping (CDLFM) model. They obtain user rating behavior from the source domain and perform matrix factorization to alleviate the cold-start problem in the target domain. The CDLFM is modeled using gradient boosted trees and multi-layer perceptron, considering feature mapping of users within and across neighborhoods.

The methods that address the cold-start problem by introducing auxiliary data or cross-domain data are primarily aimed at solving the cold-start issue for new users. However, for new item cold-start problems, data-driven strategies have not provided many solutions.

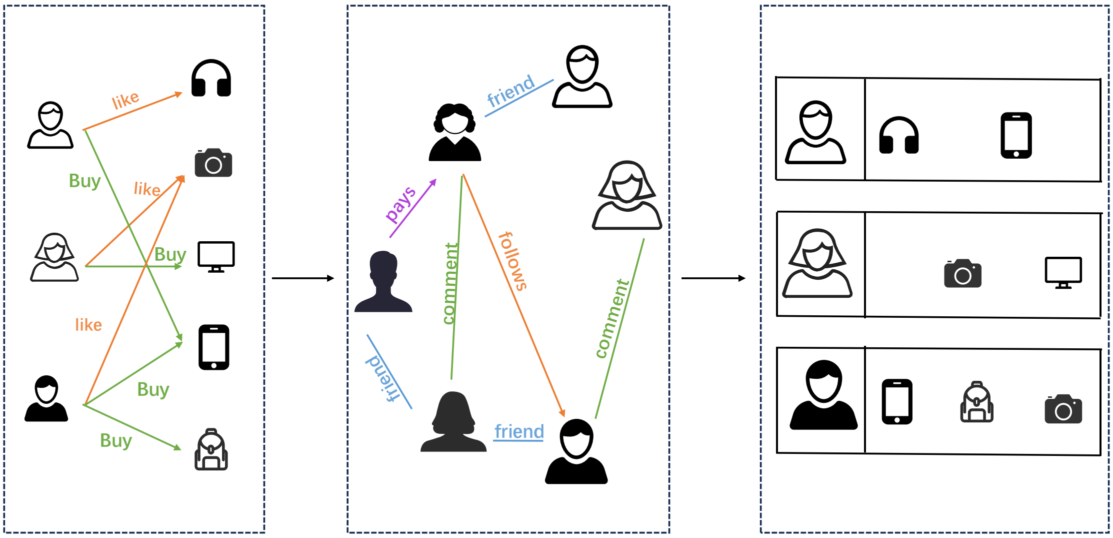

Figure 1. Framework of recommender systems based on social networks

4.2. Approach-driven Strategies

Approach-driven recommendation systems can be categorized based on the algorithms used: meta learning based approaches, session based approaches, heterogeneous graph and attributed graph based approaches, and novel approaches in recommendation systems.

4.2.1. Meta Learning Based Approaches

Meta-learning is an approach within the field of machine learning that seeks to endow algorithms with the ability to learn from past experiences how to learn. The objective is to enable rapid adaptation to new tasks from a limited number of samples, while also achieving strong generalization performance.

STRUB proposed a method that integrates meta-path embedding with word embedding to address the cold-start problem in recommendation systems. By leveraging product metadata and product names, in conjunction with existing ratings and user information, the approach trains embedded vectors associated with items. This enables the recommendation of new products by utilizing the learned embeddings to infer user preferences.

LIU utilizes an explicit preference graph to model the interactions between users and items, and employs meta-learning to infer the preferences of new users and the characteristics of new items, thereby enabling cold-start recommendations.

Kim proposed a novel gradient-based meta-learning framework for sequential recommendation, termed Meta-learning with Adaptive Weighted Loss (MELO). This framework captures the imbalance in user rating distributions and computes adaptive loss for user-specific learning accordingly.

Wang introduced a Knowledge Graph Attention Network (KGAT) that models high-order connectivity in collaborative knowledge graphs explicitly in an end-to-end fashion. It recursively propagates embeddings of node neighbors and employs an attention mechanism to differentiate the importance of neighbors. The formulation of the attention mechanism is as follows:

$$a_{i,j}^{k} = \frac{\exp(e_{i,j}^{k})}{\sum_{l \in N_i^k} \exp(e_{i,l}^{k})}$$

Where i and j denote the interacting entities, and k represents different relationship models. $e_{i,j}^{k}$ represents the attention weight between entities i and j with relation k, which is calculated as follows:

$$e_{i,j}^{k} = f(W_2^k \sigma(W_1^k [h_i^k || h_j^k]))$$

However, this method encounters certain bottlenecks in the representation of nodes and edges. Due to the ever-increasing number of nodes and diverse types of relationships, the knowledge graph requires constant updates and maintenance, presenting new challenges for recommendation systems.

4.2.2. Session Based Approaches

Session-based recommendation algorithms are designed to provide personalized recommendations by leveraging contextual information such as the user's current environment, behavior, and preferences, as well as the interests and needs demonstrated by the user's consecutive sessions.

Song utilized recurrent neural networks(RNN) to model dynamic user behavior and graph attention neural networks to model context-dependent social influence, dynamically inferring influencers based on the user's current interests. However, this approach may result in the filtering out of important but less frequently used items, which can negatively impact the recommendation effectiveness of the system.

XU proposed a Graph Contextualized Self-Attention Network (GC-SAN), which incorporates graph neural networks and self-attention mechanisms, designed for session-based recommendation tasks.

CHEN proposed a lossless encoding scheme and an edge-order-preserving aggregation layer based on Gated Recurrent Units (GRU). To address the inefficiency in capturing long-range dependencies, a fast graph attention layer was introduced, which effectively captures long-range dependencies by propagating shortcut connection information. However, this method is only suitable for shorter sequences and cannot handle long sequences.

Ur Rehman introduced a model termed Contextually Augmented Meta-Learning (CAML), which enhances contextual information by forwarding user context to a Data Augmentation Unit (DAU). The Meta-Learner (MetaL) is associated with the DAU to learn contextual preferences and generate relevant recommendations or rating predictions.

WANG introduced a novel approach called Global Context Enhanced Graph Neural Networks (GCE-GNN), which leverages item transitions across all sessions to infer user preferences within the current session. However, for certain datasets, the global feature vectors do not significantly contribute to the improvement of recommendation performance.

Table 2. Comparison of advantages and disadvantages of method-driven cold start solution algorithms

Strategy	Taxonomy	Algorithm Name	Advantages	Disadvantages
meta learning	New Item	MetaKG	1.Effective collection of advanced writing tasks	Prone to noise influence.
			2. Adaptive task selection	
	New User	ML2E	1. The training process necessitates only a small amount of first-order gradient information.	1. Requires prior knowledge.
			2. Robust generalization capabilities across diverse tasks.	2. Gradient-based training involves high complexity.
		KGAT	End-to-end utilization of higher-order information.	Lacks explicit modeling of user interests in relationships.
session	New Item	DSMR	1. r temporal dynamics with strong scalability.	Poor interpretability of recommendations.
			2. The incorporation of item description documents enhances the precision of recommendations.	
		GC-SAN	1. Graph structures are capable of capturing the complex associations between user behaviors, providing richer recommendation information.	1. High computational complexity.
			2. Dynamically learning the degree of association between nodes to enhance accuracy.	2. Lack of historical data precludes
	New User	GCE-GNN	1. Fusing user behavior sequences, global context information, and item features aids in enhancing the accuracy and diversity of recommendations.	1. High computational complexity.
			2. Graph neural networks effectively capture the associations and dependencies between user behaviors.	2. Efficiency decreases on large-scale datasets.
		CAML	1. Utilizing context features as input results in more accurate	1. Requires a lot of historical data for model training.
			2. Simultaneously handling multiple recommendation tasks and achieving superior performance on each.	2. Significant computational and storage resource.
Heterogeneous graph and attributed graph	New Item	AGNN	1. Utilizing attribute graphs can address the complete cold-start problem.	1. Significant computational and storage resource
			2. Gated systems aggregate various domain-specific information.	2. Susceptible to the influence of noise.
	New User	GCN4RS	Integrating heterogeneous graph information with user behavior data to enhance recommendation accuracy.	1. Data sparsity.
				2. Requires substantial computational resources and has a longer training time.
		IGNN	1. Enhancing model capacity through gated attention structures.	Poor interpretability of recommendations.
			2. Addressing the absence of user preference embeddings.	

4.2.3. Heterogeneous Graph and Attributed Graph Based Approaches

Heterogeneous graphs encompass a graphical structure where multiple types of nodes and edges coexist, each possessing distinct semantic meanings and attributes. Within such graphs, nodes can represent various entities or concepts, while edges denote different types of relationships or connections. An attributed graph, on the other hand, refers to a graph in which each node and edge is endowed with specific attributes or features. These attributes provide additional information about the nodes and edges, such as the category of a node, the attribute values of a node, or the weight of an edge. Through the analysis of heterogeneous and attributed graphs, a more comprehensive understanding of user preferences can be attained.

QIAN developed a novel framework called Attribute Graph Neural Networks (AGNN) by leveraging attributed graphs. This approach is capable of generating preference embeddings for cold users or items by learning the distribution of attributes using an extended variational autoencoder structure. However, this method relies on attribute information provided by external knowledge bases, which may lead to loss and inaccuracy in the results [42].

Wu proposed a personalized course recommendation model for predicting course completion. This model constructs an online course learning heterogeneous graph for students' courses, employing graph neural networks to generate embeddings for course nodes within the heterogeneous graph. Subsequently, it integrates learning status representations and course embeddings through an interaction mechanism to predict the completion likelihood of the student's next course.

Ge proposed a novel recommendation algorithm based on graph convolutional networks (GCN4RS), which employs two sets of graph convolution operations to simultaneously utilize two different types of interactive information. By integrating both heterogeneous graph information and user historical behavior sequence information, the algorithm enhances recommendation performance.

Gao designed and constructed a method called the Infographic Neural Network (IGNN), which incorporates user features and item attributes into an information graph. They also provided an improved variational graph autoencoder to address the issue of preference reconstruction embedding.

4.2.4. Novel Approach in Recommendation Systems

Researchers have also tried to improve the approaches and bring about novelty in the methods to overcome cold start problems.

Pan proposed a novel similarity model, the Popularity Mean Squared Deviation (PMSD) model, which takes into account the influence of popular items, the average difference in ratings between two users on common items, and the non-numeric information of ratings. This strategy considers the deviation between two popular items among two users and uses the similarity of the common ratings between two users as a weight to adjust the deviation.

Zhang proposed a strategy that engages in mutual relation mining to extract binary relationships between item attributes and utilizes these interrelated attributes to define the commonalities between items and new items, thereby addressing the new item cold-start problem.

ZHAO proposed a graph neural network completion model that integrates high-order semantic features of two target entities on top of the graph neural network. Additionally, the model introduces a multi-task learning mechanism, utilizing knowledge graph embeddings to provide auxiliary information for the recommendation module, achieving high-level information interaction and enhancing the model's generalization capability. However, the use of multi-task learning may pose a risk of overfitting.

MEI proposed a group preference fusion strategy that integrates a dual-layer attention mechanism, merging user preferences and group preferences vectors. This approach enhances focus through the dual-layer attention mechanism, aiming to improve the recommendation accuracy and interpretability of the model. While the model demonstrates high scalability and practicality in real-world applications, it suffers from high computational complexity, and the results may lack interpretability.

5. Discussion and Directions for Future Research

Recommendation systems typically face two key issues that affect recommendation accuracy: (1) data sparsity and (2) cold start. Data sparsity arises from a lack of item ratings, while the cold start problem is caused by new users or items that cannot be analyzed by the recommendation system. The cold start problem is the most challenging of these two issues as it affects the interaction between new users and the recommendation system. The cold start issue arises due to the presence of new users or new items in the recommendation dataset. These new users or items pose recommendation problems with respect to rating data, and new user interaction data is unavailable. Many researchers have attempted to develop several approaches and algorithms to cater to these problems. With Web 3.0, recommendation systems now employ hybrid ensemble algorithms, utilizing location information and a vast amount of other data collected from devices and sensors. These devices and sensors provide real-time health information, dietary habits, and data on daily life and routines. Advances in computing power, as well as the development of machine learning and deep learning algorithms, have also contributed to this endeavor.

Therefore, this study has effectively compiled a multitude of different recommendation system methods that have been developed and experimentally validated to produce better results on the new user and new item cold start problems. This review systematically analyzes the methods and algorithms that researchers have used for the cold start problem, the benchmark algorithms they have used to compare the effectiveness of these methods, and the metrics that authors have employed to evaluate algorithm efficiency. The study finds that basic collaborative filtering algorithms such as user-based collaborative filtering and item-based collaborative filtering are widely used as benchmark algorithms by many researchers. However, researchers have also employed a variety of algorithms without noticing a common pattern in this study. In terms of metrics for evaluating the effectiveness of the proposed algorithms and methods, Mean Absolute Error (MAE), precision, and recall are the most widely used by different researchers, followed by Root Mean Square Error (RMSE). This provides an opportunity for future recommendation system scholars to study and report on which key benchmark algorithms and metrics should be considered to evaluate the performance of algorithms in cold start scenarios.

Future research directions will focus on leveraging deep learning and representation learning techniques to handle multi-modal data and complex relationships, such as enhancing user and item representations through graph neural networks. Additionally, meta-learning and online learning will be employed to address the cold-start problem, enabling rapid adaptation to new users or items. Furthermore, research emphasis will be placed on enhancing the interpretability and transparency of recommendation systems, ensuring fairness and diversity, as well as strengthening privacy protection and data security. Cross-domain and cross-modal recommendation systems, as well as optimization of user engagement and feedback loops, will further advance the field. With technological advancements and the diversification of user needs, solutions to the cold-start problem will continue to evolve, becoming a key driving force for ongoing innovation in the research and application of recommender systems.

6. Conclusion

In summary, this review presents a collection of different methods and algorithms that recommendation system researchers have adopted over the past decade to address the new user and new item cold start problems. The main contribution of this study is an exhaustive research and synthesis of 52 articles on recommendation systems that propose methods and algorithms to mitigate the cold start problem. The study effectively synthesizes various strategies into two broad categories: data-driven strategies and method-driven strategies. It also collects details such as the domains of recommendation system research, benchmark algorithms used for comparison, and metrics used for evaluation. Therefore, this study can serve as a comprehensive guide for future recommendation system researchers to understand the research done so far on the cold start problem.

Acknowledge

This study was funded by the Chongqing University of Technology's High-Quality Development Project for Postgraduate Education (Grant No. gzlcx20233256).

References

[1] Richa, BP. Combining trust and reputation as user influence in cross domain group recommender system (cdgrs). [J]. Journal of Intelligent and Fuzzy Systems, 2020(38): 6235–6246.

[2] Viktoratos, I, Tsadiras, etal. Combining community-based knowledge with asso- ciation rule mining to alleviate the cold start problem in context-aware recommender systems. [J]. Expert Systems with Applications, 2018(11):78-90.

[3] Chen, L., & Pu, P. Critiquing-based recommenders: Survey and emerging trends. [J]. User Modeling and User-Adapted Interaction, 2012(22), 125–150

[4] Puthiya Parambath, S. A., & Chawla, S. (2020). Simple and effective neural-free soft-cluster embeddings for item cold-start recommendations. Data Mining and Knowledge Discovery, 34, 1560–1588.

[5] Rohani, V. A., Kasirun, Z. M., Kumar, S., & Shamshirband, S. (2014). An effective recommender algorithm for cold-start problem in academic social networks. Mathematical Problems in Engineering, 2014.

[6] Zhang, Q., Wu, D., Lu, J., Liu, F., & Zhang, G. (2017). A cross-domain recommender system with consistent information transfer. Decision Support Systems, 104, 49–63.

[7] Liu, J. H., Zhou, T., Zhang, Z. K., Yang, Z., Liu, C., & Li, W.M. (2014b). Promoting cold-start items in recommender systems. PLoS ONE, 9, 1–14.

[8] MAO Qian, XIE Weicheng, QIAO Yitian, HUANG Xiaolong, DONG Gang. Survey on Solving Cold Start Problem in Recommendation Systems[J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(5): 1197-1210.

[9] Suryana, N., & Basari, A. S. B. H. (2018). An understanding and approach solution for cold start problem associated with recommender system: A literature review. Journal of Theoretical and Applied Information Technology, 15,9.

[10] Camacho, L. A. G., & Alves-Souza, S. N. (2018). Social network data to alleviate cold-start in recommender system: A systematic review. Information Processing and Management, 54, 529–544.

[11] CHANG Hao, WU Wei, SHU Xiaodong, et al. Application of content-based recommendation system in homestay[C]//2021 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS):2021:493-496.

[12] JEEVAMOL J, RENUMOL V. An ontology-based hybrid e-learning content recommender system for alleviating the cold-start problem[J]. Education and Information Technologies, 2021, 26: 4993-5022

[13] CRUZ AF T.CORONEL A D. Towards developing a content-based recommendation system for classical music [C] // Information Science and Applications: ICISA 2019. Springer Singapore. 2020:451-462.

[14] Hongcan Yan, Ziru Wang, Weifang Li. Fuzzy Clustering Collaborative Filtering Recommendation Algorithm with Time Evolution [J]. Computer Engineering and Science, 2021, 43(11): 2084-2090.

[15] ZHANG Y, SHI Z, ZUO W, et al. Joint personalized Markov chains with social network embedding for cold- start recommendation[J]. Neurocomputing, 2020, 386: 208-220.

[16] FAN Y Q, DU Y J, CHENG L J, et al. An improved collaborative filtering recommendation method with user trust preference fusion[J]. Journal of Xihua University (Natural Science), 2015, 34(4): 8-12.

[17] WU QCHENG XSUN E, et al. Collaborative filtering algorithm based on optimized clustering and fusion of user attribute features[C]//2021 4th international Conference on Data Science and information technology.2021:125-12.

[18] XUE F, HE X, WANG X, et al. Deep item-based collaborative filtering for top- n recommendation[J]. ACM Transactions on Information Systems, 2019, 37(3): 1-25.

[19] ESTEBAN AZAFRAAROMERO C. Helping university students to choose elective courses by using a hybrid multicriteria recommendation system with genetic optimization [J]. Knowledge-Based Systems 2020, 194 (12): 105385.

[20] WU Z Y,LI M,TANG Y, et al. Exercise recommendation based on knowledge concept prediction [J].Knowledge Based Systems, 2020, 210:106481.

[21] HU J, LIU L, ZHANG C, et al. Hybrid recommendation algorithm based on latent factor model and Personal Rank[J]. Journal of Internet Technology, 2018, 19(3): 919-926.

[22] RAIGOZA J, KARANDE V. A study and implementation of a movie recommendation system in a cloud- based environment[J]. International Journal of Grid and High Performance Computing, 2017, 9(1): 25-36.

[23] LIKA B, KOLOMVATSOS K, HADJIEFTHYMIADES S. Facing the cold start problem in recommender systems[J]. Expert Systems with Applications, 2014, 41(4): 2065-2073.

[24] MU R, ZENG X. Auxiliary stacked denoising autoencoder based collaborative filtering recommendation[J]. KSII Transactions on Internet and Information Systems, 2020, 14(6): 2310-2332.

[25] XU C. A novel recommendation method based on social network using matrix factorization technique[J]. Information Processing & Management, 2018, 54(3): 463-474.

[26] Ahmadian, S., Afsharchi, M., & Meghdadi, M. (2019). An effective social recommendation method based on user reputation model and rating profile enhancement. Journal of Information Science, 45, 607–642.

[27] Ojagh, S., Malek, M. R., & Saeedi, S. (2020). A social-aware recommender system based on user's personal smart devices. ISPRS International Journal ofGeo-Information,9.

[28] Herce-Zelaya, J., Porcel, C., Bernab´e-Moreno, J., Tejeda-Lorente, A., & Herrera-Viedma, E. (2020). New technique to alleviate the cold start problem in recommender systems using information from social media and random decision forests. Information Sciences, 536, 156–170.

[29] Ghavipour, M., & Meybodi, M. R. (2019). Stochastic trust network enriched by similarity relations to enhance trust-aware recommendations. Applied Intelligence, 49, 435–448.

[30] REN H, LIU B S, SUN J Y. Advances and perspectives on knowledge transfer based cross- domain recommendation[J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(11): 1813-1827.

[31] WANG X, PENG Z, WANG S, et al. CDLFM: cross-domain recommendation for cold-start users via latent feature map ping[J]. Knowledge and Information Systems, 2020, 62: 1723-1750.

[32] Wang, X., Peng, Z., Wang, S., Yu, P. S., Fu, W., Xu, X., & Hong, X. (2020). Cdlfm: cross-domain recommendation for cold-start users via latent feature mapping. Knowledge and Information Systems, 62,1723–1750.

[33] STRUB F, MARY J, GAUDEL R. Hybrid collaborative filtering with autoencoders[J].

[34] LIU H, WANG L, LI P, et al. Relation- propagation meta learning on an explicit preference graph for cold- start recommendation[J]. Knowledge- Based Systems, 2023, 272: 110579.

[35] KIM M, YANG Y, RYU J H, et al. Meta-learning with adaptive weighted loss for imbalanced cold- start recommendation[J]. arXiv:2302.14640, 2023.

[36] WANG X, HE X, CAO Y, et al. KGAT: knowledge graph attention network for recommendation[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2019.

[37] SONG W, XIAO Z, WANG Y, et al. Session- based social recommendation via dynamic graph attention networks[C]// Proceedings of the 12th ACM International Conference on Web Search and Data Mining. New York: ACM, 2019: 555-563.

[38] XU C, ZHAO P, LIU Y, et al. Graph contextualized self-attention network for session- based recommendation[C]// Proceedings of the 28th International Joint Conference on Artificial Intelligence, Macao, China, Aug 10-16, 2019: 3940-3946.

[39] CHEN T, WONG R C W. Handling information loss of graph neural networks for session- based recommendation [C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2020: 1172-1180.

[40] UR REHMAN I, ALI W, JAN Z, et al. CAML: contextual augmented meta- learning for cold- start recommendation[J]. Neurocomputing, 2023, 533: 178-190.

[41] WANG Z, WEI W, CONG G, et al. Global context enhanced graph neural networks for session-based recommendation[C]//Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2020: 169-178.

[42] QIAN T, LIANG Y, LI Q, et al. Attribute graph neural networks for strict cold start recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2020, 34(8): 3597-3610.

[43] WU Z Y, ZHANG G T, HUANG L, et al. Course recommendation based on heterogeneous graph embedding and session interaction[J/OL]. Computer Engineering [2023-07- 20].

[44] GE Y, CHEN S C. Graph convolutional network for recommender systems[J]. Journal of Software, 2020, 31(4): 1101- 1112

[45] GAO W, ZHU F L, LI D Z, et al. Realization of graph neural network in cold start recommendation[J]. Computer Engineering and Design, 2022, 43(9): 2557-2566.

[46] Pan, R., Ge, C., Zhang, L., Zhao, W., & Shao, X. (2020). A new similarity model based on collaborative filtering for new user cold start recommendation. IEICE Transactions on Information and Systems E103D,1388–1394.

[47] Zhang, Z. P., Kudo, Y., Murai, T., & Ren, Y.G. (2019). Addressing complete new item cold-start recommendation: A niche item-based collaborative filtering via interrelationship mining. Applied Sciences (Switzerland),9.

[48] ZHAO W, LI Y, FAN T, et al. A novel embedding learning framework for relation completion and recommendation based on graph neural network and multi- task learning[J/OL]. Soft Computing [2023-07-20]. https://doi.org/10.1007/s00500-021-06617-0.

[49] MEI Y Z, HU Z L, ZHU X J. Research on group preference fusion strategy based on two-layer attention mechanism[J]. Computer Engineering and Applications, 2023, 59(9): 272-279.

[50] Richa, BP. (2020). Combining trust and reputation as user influence in cross domain grouprecommender system (cdgrs). Journal of Intelligent and Fuzzy Systems, 38, 6235–6246.

[51] Rohani, V. A., Kasirun, Z. M., Kumar, S., & Shamshirband, S. (2014). An effective recommender algorithm for cold-start problem in academic social networks. Mathematical Problems in Engineering, 2014.

[52] Bobadilla, J., Ortega, F., Hernando, A., & Guti´errez, A. (2013). Recommender systems survey. Knowledge-Based Systems, 46, 109–132.