ITM Web of Conferences **78**, 04022 (2025) https://doi.org/10.1051/itmconf/20257804022
CSEIT 2025

A Knowledge Graph Recommendation Model Based on Reinforcement Learning and Comparative Learning

Xianrong Zhang

School Of Computer Science & Technology, Soochow University, Suzhou, 215000, China

Abstract. Recommendation system has become the core technology of recommendation platforms, with the continuous growth of user demand, how to improve the personalization and explanation quality of recommendation system has emerged as a significant area of research. This paper proposes a dynamic recommendation optimization method based on reinforcement learning and comparative learning which aims to solve the problems of insufficient flexibility of static weighting strategies and low quality of recommendation explanation in traditional recommendation systems. The method combines knowledge graph, optimizes the embedded representations of users and items through comparative learning, and introduces the Q-Learning algorithm to dynamically select positive-negative sample pairs and knowledge graph paths in order to improve recommendation accuracy and explanation transparency. The experimental findings demonstrate that this method considerably surpasses static optimization techniques and currently available knowledge graph-based recommendation models. After 10 rounds of training, the Accuracy (AUC) and F1 scores of the test set are both improved by about 3% over the existing models, and the Recall@K value is also higher than other models. In this paper, an adaptive recommendation framework integrating reinforcement learning and comparative learning is constructed, which provides theoretical support and practical foundation for more efficient personalized recommendations and diversified explanations.

1 Introduction

Recommender systems have become an indispensable part of daily online services by analyzing user behavior to predict users' interests and thus provide personalized recommendations. However, most of the traditional recommender systems rely on the interaction data between users and items, and this approach suffers from the problems of sparse data and cold start. In this context, Knowledge Graph (KG) is widely used as a kind of external information in recommender systems. By organizing entities and their relationships, Knowledge Graph can provide rich background information for recommender

2227405072@stu.suda.edu.cn

systems, thus helping to alleviate the data sparsity problem and improve the accuracy of recommendation results.

Despite the remarkable progress of knowledge graph-based recommender systems, however, they still face the challenges of static optimization strategies leading to insufficient response to changes in user interests, insufficient explanatory nature of the recommender system, and failure of predefined explanations to reflect user behavior in real time.

Currently, scholars have made various attempts in the field of knowledge graph-assisted recommender systems. For example, the Knowledge Graph Attention Network (KGAT) model explicitly models higher-order relationships in the knowledge graph through graph attention networks, thereby improving recommendation accuracy and interpretability [1]. In addition, the Explainable Method for Knowledge Aware Recommendation (Ekar) approach provides more explanatory recommendations by modeling the recommendation problem as a sequential decision-making process through reinforcement learning that dynamically selects the paths from the user to the items [2]. The RippleNet approach, on the other hand, utilizes graph attention mechanism to propagate the user's interest to generate interpretable recommendation paths [3].

The Large Language Model (LLM) and KG for Personalized News Recommendation framework (LKPNR) approach enhances news semantic representation and mitigates the long-tail problem by introducing large-scale language models and knowledge graphs. The approach utilizes LLM to understand text semantics and combines KG to mine higher-order structural information among entities, thus enhancing the accuracy of personalized news recommendation [4]. The generic Recommendation Model enriches the representation of users and items and improves the prediction accuracy by introducing a knowledge graph as additional information into the recommendation system. It supports the construction of KG by extracting information from different data sources and using it to train recommendation models to enhance the generalization ability of the system [5].

However, most of the existing reinforcement learning methods focus on path selection and recommendation optimization, but they usually rely on static strategies for path selection, failing to take full advantage of the potential of contrast learning in optimizing sample pair selection and path navigation. Contrastive learning has been shown to be very effective in generating effective user and item embedding representations, but there is still a large room for improvement in the application of dynamic optimization for sample selection and path navigation.

To overcome these problems, this paper proposes a new framework for recommender systems, Knowledge Graphs and Interaction Contrastive Learning with Reinforcement Learning (KGIC-RL). The method combines reinforcement learning and contrast learning and dynamically adjusts the selection of recommendation paths and sample pairs through the Q-Learning algorithm to optimize recommendation accuracy and explanation transparency. Under the Q-Learning framework, KGIC-RL not only optimizes the sample selection based on the user's historical interactions but also guides the dynamic adjustment of the paths to better adapt to the changes in user interests. Experimental results show that KGIC-RL significantly outperforms existing methods in both recommendation accuracy and explanation quality, providing a more flexible and efficient recommendation optimization scheme.

2 Data and methodology

2.1 Data sources

In this paper, three publicly available benchmark datasets for recommender systems, Book-Crossing, MovieLens-1Million (MovieLens-1M) and Last. Frequency Modulation (Last.FM), is used to evaluate the performance of the proposed model KGIC-RL. These three datasets have different sizes, structures and sparsities, which make the experimental results more general and convincing. The basic statistics of each dataset are shown in Table 1.

Table 1 Statistical data for the pre-processed dataset in this study

		Book-Crossing	MovieLens-1M	Last.FM
User-item Interaction	# users	17,860	6,036	1,872
	# items	14,967	2,445	3,846
	#interactions	139,746	753,772	42,346
Knowledge Graph	# entities	77,903	182,011	9,366
	# relations	25	12	60
	# triplets	151,500	1,241,996	15,518

The Book-Crossing dataset is derived from the Book-Crossing community and contains explicit user ratings of books ranging from 0 to 10. The source of this dataset is http://www2.informatik.uni-freiburg.de/~cziegler/BX/. It covers a broader range of book categories and user preferences. The MovieLens-1M dataset contains about 1 million user ratings of movies ranging from 1 to 5 and is one of the most widely used standard datasets in recommender systems. The source of this dataset is https://grouplens.org/datasets/movielens/1m/. Last.FM is collected from Last.FM online music systems, containing listening history with around 2 thousand users. It comes from https://grouplens.org/datasets/hetrec-2011/.

In order to unify the recommendation learning paradigm, this paper refers to Knowledge graph convolutional networks (KGCN) and RippleNet, and converts all the explicit feedback in the above datasets into binarized implicit feedbacks: i.e., the records where ratings are present are regarded as positive samples (denoted as 1), and an equal number of randomly sampled records from the non-ratings or unlistened-to items are regarded as negative samples (denoted as 0)[6,7]. This treatment not only simplifies the rating modeling process but also allows the model to be trained uniformly with a click/non-click dichotomous target.

In addition, to introduce entity semantic information, this paper utilizes Microsoft Satori to construct a Sub-Knowledge Graph (sub-KG) for each dataset, which contains structural information <Head Entity, Relationship, Tail Entity> in ternary format. The sub-KG is a high-confidence (confidence > 0.9) subset extracted from the whole Knowledge Graph, which matches the entity IDs (Satori IDs) based on the item names to realize the alignment of the items (e.g., books, movies, musicians) with the Knowledge Graph entities. The aligned item IDs are matched with the head entities in the triples, and all the highly correlated triples are extracted to form the sub-KGs. all the triples are converted to a uniform numeric number and saved in kg_final.txt format for model input.

In terms of data preprocessing, this paper standardized and cleaned all raw data as follows. The raw ratings are converted into positive and negative samples (e.g., MovieLens ≥4 is a positive sample) by setting a threshold, and saved in a unified format as ratings_final.txt file. The user interaction dataset is split into training, validation, and testing subsets following a 6:2:2 ratio to ensure the independence of model evaluation at different stages.

Eventually, the preprocessed data is converted into a standardized training format and fed into the KGIC-RL model for recommendation learning. Each user not only has a history of the items they interacted with but also associates entity paths related to their interests, enabling graph-based recommendation and interpretation.

2.2 Methodology

2.2.1 Concepts and principles

KG is a graphical representation of structured semantic information that integrates side-by-side information with a collaborative filtering model to solve possible cold-start problems [8,9]. It describes the diverse relationships between entities in the real world through a ternary form (entity-relationship-entity). In recommendation scenarios, knowledge graphs can establish connections between items of interest to users and other semantically related entities (e.g., authors, directors, tags, categories, etc.), to mine deep interest paths and contextual information, and to assist the model in making more accurate and reasonable recommendations.

Contrastive learning is a self-supervised learning method whose core idea is to acquire more distinct and discriminative feature representations by constructing pairs of positive and negative samples, maximizing the similarity between positive samples, and minimizing the similarity between negative samples. In KGIC-RL, this study designs both intra-graph and inter-graph comparison strategies: intra-graph comparison is used to enhance the consistency of entity representations in the graph structure, while inter-graph comparison is used to improve the stability of user preference expressions under different perspectives. Through the dual contrast mechanism, the model can capture the semantic relationships between entities more effectively and improve the generalization ability and robustness of recommendations.

The KGIC-RL model combines knowledge graph and reinforcement learning, aiming to optimize recommendation effectiveness through comparative learning while optimizing path selection strategies through reinforcement learning. The core idea is to model the recommender system as a reinforcement learning task, where each state is represented as the historical interaction information between the user and the recommended item, and the actions are the path selection and optimization model parameters. Based on the user's historical interaction and the entity relationship information in the knowledge graph, the system gradually improves the recommendation strategy by continuously optimizing the Q-value, so as to obtain better recommendation accuracy and interpretability.

The model itself also combines graph neural networks and contrast learning to further enhance the embedding learning capability of entities and relationships in the knowledge graph. By introducing contrast learning loss, KGIC-RL is able to optimize the selection of positive and negative sample pairs and learn more accurate representations of user preferences and item features.

2.2.2 Model characteristics

In traditional recommender systems, the model usually selects recommendation paths based on a fixed policy. In contrast, in KGIC-RL, reinforcement learning (especially Q-Learning) is used to dynamically optimize path selection and hyperparameter tuning, which makes the model more adaptable in the face of changing user preferences. Through the reward mechanism, the model is able to effectively explore and utilize the relational information in the knowledge graph in the recommendation task.

KGIC-RL enhances the accuracy and interpretability of recommendations through the rich semantic information provided by the knowledge graph. The entities and their relationships in the knowledge graph can help the model understand the potential connections between users and items, providing additional contextual information for the recommender system.

In order to demonstrate the effectiveness of the model proposed in this paper, KGIC-RL is compared with state-of-the-art methods including: the Collaborative Filtering (CF) based method Bayesian personalized ranking from implicit feedback (BPRMF), the embedding-based method RippleNet, the path-based method Personalized entity recommendation (PER) and the Graph Neural Network (GNN) based methods KGCN, Knowledge graph attention network for recommendation (KGAT), Knowledge Graph for interactions (KGIN)[6,7,10,11,12,13].

2.2.3 Experimental procedure and parameters

The main steps of the experimental process include data loading, model training and evaluation. During model training, the knowledge graph paths of users and items are dynamically adjusted by contrast learning and reinforcement learning algorithms. In each training cycle (epoch), the system selects appropriate hyper-parameters (e.g., intra/inter-graph contrast loss weights) based on the current model state (e.g., Q-value table) and updates the parameters based on the model's performance on the validation set. The experimental parameters are shown in Table 2.

Table 2 Experimental parameters

learning rate	batch size	epochs	n_layer	lambda_intra & lambda_inter
0.04	2048	10	2	1e-6

In this paper, the learning rate is set to be 0.04, epochs to be 10, batch size to be 2048, for the neural network, the layer is configured to be 2, and the regularization term coefficient is set to be 1e-6 to avoid overfitting.

The dataset is first loaded and transformed through the data preprocessing module, and the processing includes binarization of the scoring data, ternary construction of the knowledge graph, and division of the training, validation, and test sets.

In each training cycle, the model optimizes the recommendation results by calculating the loss of positive and negative sample pairs and adjusts the hyperparameters and path selection strategies through the Q-Learning mechanism of reinforcement learning. After every 5 training cycles, the model evaluates its effectiveness when evaluated on the validation set and further optimizes the hyperparameters through the reward mechanism of Q-Learning.

Following the training phase, the model's performance is assessed on the test set using AUC and F1 scores to evaluate its recommendation precision and recall capability.

With the above approach, KGIC-RL is able to provide interpretable recommendation paths while enhancing recommendation accuracy, thus improving user experience and system transparency.

2.3 Assessment of Indicators

To comprehensively assess the proposed methodology, this study conducts evaluations under two experimental settings: (1) For Click-Through Rate (CTR) prediction, the commonly adopted metrics AUC and F1 score are utilized. (2) In the context of top-K recommendation, Recall@K is employed to measure the effectiveness of the recommended items, with K values consistently chosen as 5, 10, 20, and 50.

3 Analysis of results

The results obtained from this study and the comparison model on different datasets are presented in Table 3.

Table 3 Comparison of experimental results

Model	Book-Crossing AUC	Book-Crossing F1	MovieLens-1M AUC	MovieLens-1M F1	Last. FM AUC	Last. FM F1
BPRMF	0.658	0.611	0.892	0.792	0.756	0.701
RippleNet	0.721	0.647	0.919	0.842	0.776	0.702
PER	0.604	0.572	0.712	0.667	0.641	0.603
KGCN	0.684	0.631	0.909	0.836	0.802	0.708
KGAT	0.731	0.654	0.914	0.844	0.829	0.742
KGIN	0.727	0.661	0.919	0.844	0.848	0.760
KGIC-RL	0.740	0.680	0.939	0.850	0.854	0.772

On the Book-Crossing dataset, the AUC of KGIC-RL improves by 1.9% compared to KGIN, and the F1 of KGIC-RL improves by 0.9% compared to KGAT. This indicates that KGIC-RL is more capable of mining the interest paths in the sparse book ratings.

On the MovieLens-1M dataset, the AUC of KGIC-RL reaches 93.9% and the F1 reaches 85.0%, which are 2.0% and 0.6% higher than that of KGIN, respectively, reflecting the stronger feature expression ability in large-scale explicit scoring data.

KGIC-RL also achieves optimal results in the Last.FM music dataset, especially outperforming KGIN by 0.6% in the F1 metric, indicating that the method is better balanced in handling implicit feedback.

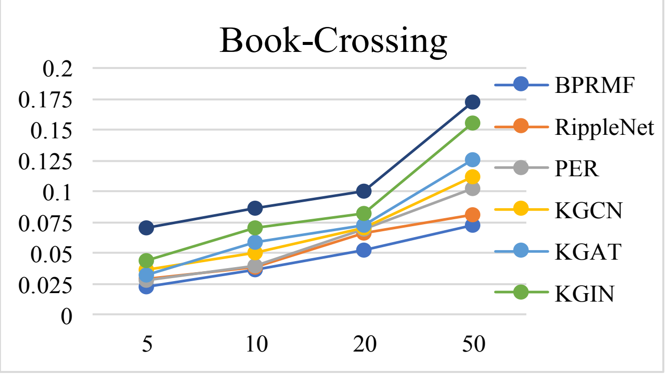

Figure 1 Recall@K Results for Book-Crossing (Photo/Picture credit : Original)

As can be seen from Figure 1, KGIC-RL significantly outperforms other models on the Book-Crossing dataset. At K=50, the Recall improvement is especially obvious, reaching about 0.175, which is more than 100% compared with BPRMF, and is also significantly better than 0.155 of KGIN (the second best model), which is leading at all K values, reflecting its strong ability to capture user interest paths in sparse book rating scenarios.

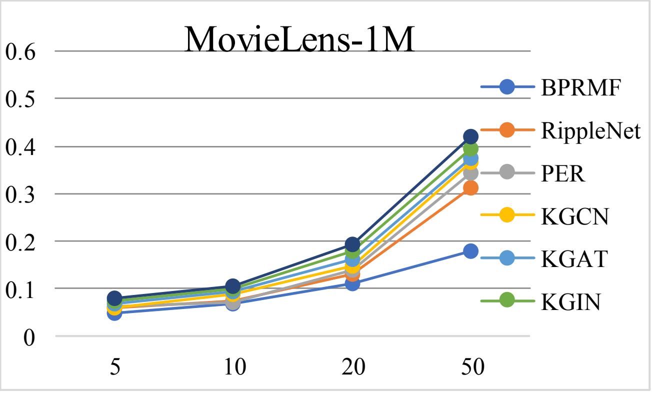

Figure 2 Recall@K Results for MovieLens-1M (Photo/Picture credit : Original)

Figure 2 demonstrates that on the MovieLens-1M dataset, KGIC-RL still achieves the optimal performance under all K values, especially at K=50 where Recall reaches 0.42+, which is slightly higher than that of KGIN and KGAT, and exhibits robustness under large-scale explicit scoring. This improvement is attributed to the fact that KGIC-RL dynamically adjusts the interest paths through reinforcement learning in large-scale explicit scoring data, which enhances the feature expression and discrimination ability. Notably, traditional models such as BPRMF perform significantly worse than graph-based models in this scenario, reflecting the limitations of pure collaborative filtering approaches.

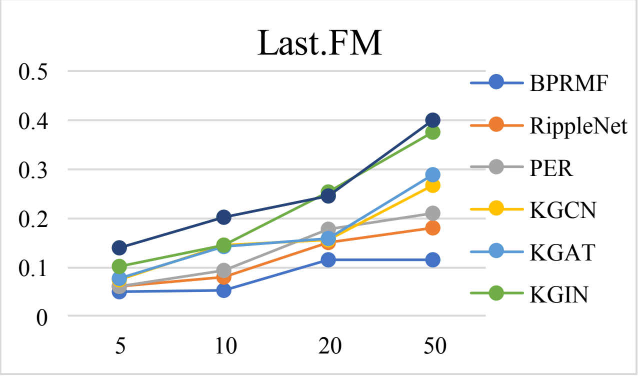

Figure 3 Recall@K results for Last. FM (Photo/Picture credit : Original)

Figure 3 illustrates that in the Last.FM (implicit feedback) dataset, KGIC-RL also maintains a steady lead, reaching Recall@5 ≈ 0.14 at K=5 and 0.40 at K=50. Among them, the curves of KGIC-RL and KGIN are higher than the rest of the models, but KGIC-RL consistently leads in all the points, which suggests that it is more effective in modeling the trajectory of interest in music. KGIC-RL always leads at all points, indicating that it is more effective in modeling music interest trajectories, especially in scenarios where the user behavior sequence is sparse but the graph information is rich.

The comparative analysis of AUC, F1 and Recall@K indicators shows that KGIC-RL achieves the best performance in all evaluation dimensions and datasets, which verifies the effectiveness of its dual mechanism combining comparative learning and reinforcement learning. At the same time, the Top-K recommendation experiment further confirms the practical value of the model, especially when the K is large, the performance is stronger, which is suitable for the recommendation scenario of “diversity + wide coverage” in the actual recommendation system. Compared with traditional methods (e.g., BPRMF, PER), KGIC-RL is more capable of combining the graph structure and user interests and shows stronger interpretability and generalization ability.

4 Conclusions

This study introduces a recommendation system model named KGIC-RL, which is based on combination of knowledge graph and reinforcement learning, and verifies its superiority in terms of recommendation effectiveness through comparison experiments alongside various state-of-the-art recommendation approaches (e.g., BPRMF, RippleNet, PER, KGCN, KGAT, and KGIN) on multiple public datasets. superiority in recommendation effectiveness. With the introduction of graph information and reinforcement learning adaptive path selection, KGIC-RL is able to show strong generalization ability and accuracy in both explicit and implicit feedback scenarios and especially makes significant progress in modeling complex user interest propagation paths.

Experimental results show that KGIC-RL outperforms other models in several evaluation metrics, such as AUC, F1 score, and Recall@K, especially in large-scale datasets (e.g., MovieLens-1M and Last.FM), where it demonstrates its superior recommendation ability and stability. Specifically, KGIC-RL consistently outperforms other methods in Top-K recommendation tasks with different K-values, proving its potential application in real recommendation scenarios. By optimizing the embedded representations of users and items through comparative learning and dynamically adjusting the recommendation paths with reinforcement learning, KGIC-RL achieves high recommendation accuracy and improves the interpretability of recommendations at the same time.

Future research directions can be carried out in the following aspects: first, the way of constructing and updating the knowledge graph can be further optimized, and more external knowledge sources can be combined to enhance the performance of the model on the cold-start problem. Second, KGIC-RL can be combined with deep reinforcement learning to explore more complex user behavior prediction models.

Overall, the KGIC-RL model provides a new way of thinking for knowledge graph-driven recommender systems, which not only improves the recommendation accuracy, but also enhances the interpretability of the model, and has a broad application prospect.

References

1. Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. 2019. KGAT: Knowledge Graph Attention Network for Recommendation. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19), Anchorage, AK, USA, August 4–8, 2019. ACM, New York, NY, USA, 950–958. https://doi.org/10.1145/3292500.3330989
2. Weiping Song, Zhijian Duan, Ziqing Yang, Hao Zhu, Ming Zhang, and Jian Tang. 2022. Ekar: An Explainable Method for Knowledge Aware Recommendation. arXiv preprint arXiv:1906.09506v3.
3. Giacomo Balloccu, Ludovico Boratto, Gianni Fenu, and Mirko Marras. 2022. Post Processing Recommender Systems with Knowledge Graphs for Recency, Popularity, and Diversity of Explanations. Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR’22), July

11–15, 2022, Madrid, Spain. ACM, 11 pages. https://doi.org/10.1145/3477495.3532041

4. Chen H, Xie R, Cui X, et al. LKPNR: LLM and KG for Personalized News Recommendation Framework[J]. arXiv preprint, 2023. arXiv:2308.12028.
5. Mao Y, Mokhov S A, Mudur S. Application of Knowledge Graphs to Provide Side Information for Improved Recommendation Accuracy[J]. Science of Computer Programming, 2021. arXiv:2101.03054.
6. Hongwei Wang, Miao Zhao, Xing Xie, Wenjie Li, and Minyi Guo. 2019. Knowledge graph convolutional networks for recommender systems. In WWW. 3307–3313.
7. Hongwei Wang, Fuzheng Zhang, Jialin Wang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. 2018. Ripplenet: Propagating user preferences on the knowledge graph for recommender systems. In CIKM. 417–426.
8. Weiran Pan, Wei Wei, and Xian-Ling Mao. 2021. Context-aware Entity Typing in Knowledge Graphs. In Findings of EMNLP. 1–8.
9. Ding Zou, Wei Wei, Xian-Ling Mao, Ziyang Wang, Minghui Qiu, Feida Zhu, and Xin Cao. 2022. Multi-level Cross-view Contrastive Learning for Knowledge-aware Recommender System. In SIGIR. 1358–1368.
10. Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In UAI. 452461.
11. Xiao Yu, Xiang Ren, Yizhou Sun, Quanquan Gu, Bradley Sturt, Urvashi Khandelwal, Brandon Norick, and Jiawei Han. 2014. Personalized entity recommendation: A heterogeneous information network approach. In WSDM. 283–292.
12. Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. 2019. Kgat: Knowledge graph attention network for recommendation. In SIGKDD. 950–958.
13. Xiang Wang, Tinglin Huang, Dingxian Wang, Yancheng Yuan, Zhenguang Liu, Xiangnan He, and Tat-Seng Chua. 2021. Learning Intents behind Interactions with Knowledge Graph for Recommendation. In WWW. 878–887.