2023 7th International Conference on Electrical, Mechanical and Computer Engineering (ICEMCE)

Knowledge Aggregation of Importance Sampling Based on GNN for Music Recommendation

Han Xiao
Beijing Institute of Graphic Communication, Beijing 100000, China
boboxingren@163.com

Dezhi Sun*
Beijing Institute of Graphic Communication, Beijing 100000, China
*Corresponding author: upperbound@163.com

Fucheng You
Beijing Institute of Graphic Communication, Beijing 100000, China
youfucheng@bigc.edu.cn

2023 7th International Conference on Electrical, Mechanical and Computer Engineering (ICEMCE) | 979-8-3503-8287-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICEMCE60359.2023.10490603

***Abstract*—With the development of mobile network and streaming media,it is difficult for people to find what they are really interested in from the massive music information. The recommendation system can solve this problem well. The traditional recommendation method ignores the rich information of graph structure, and GNN has advantages in graph representation learning. Therefore, this paper proposes a music recommendation model KAISG, which captures user preferences through the knowledge aggregation of important sampling based on GNN. Firstly, the item knowledge graph is constructed into a topological structure, and the correlation between items is effectively captured by mining their related attributes on KG. Then, the GNN is used to enrich the item representation and spread higher-order information by aggregating neighborhood information layer by layer. Especially, in order to aggregate the neighborhood information related to the current node more accurately, the model uses sampling based on importance to aggregate the neighborhood information of each layer when sampling the neighbors of each entity in KG, thus avoiding the introduction of noise in high-order relationships. In addition, GNN aggregates the representation and neighborhood information of a given entity, which can simulate high-order neighborhood information and capture the potential interests of users. Experiments on a real data set shows that our model is superior to the current baselines.**

Keywords-component: music recommendation; Knowledge Graph, Graph Neural Network

I. Introduction

With the rapid growth of information on the Internet, recommendation system has become a necessary tool for personalized content recommendation in various fields, such as e-commerce, social media and online entertainment. The traditional recommendation system relies on user-item interaction and collaborative filtering [1] technology to generate recommendations. However, these methods often suffer from cold start problems [2] and difficulties in capturing complex user preferences. In order to overcome these limitations, recent research focuses on improving the performance of recommendation system by using knowledge graph and Graph Neural Network (GNN) [3]. On the one hand, knowledge graph provides a structured representation of entities and their relationships, captures rich semantic information, and can improve the accuracy of recommendation. For example, in Figure 1, we can recommend *Jump then Fall* to users based on their listening to Taylor's song *Love Story,* because this song is also sung by Taylor and belongs to the same genre and album as *Love Story*. On the other hand, GNN is a powerful deep learning model, which can effectively learn the representation of nodes and edges in the graph, so as to better understand the underlying data [4].

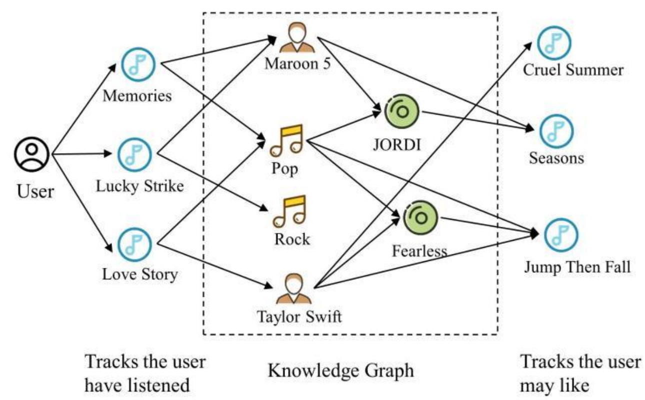

Figure 1. Illustration of knowledge graph enhanced music recommendation systems.

Recommendation methods based on knowledge graph can be roughly divided into three categories: embedding-based method, path-based method and hybrid-based method [5]. The method based on embedding uses the entities and relationships in the knowledge graph to get their embedded representations and integrate them into the recommendation framework [6]. However, this method ignores the information connectivity of knowledge graph and lacks interpretability. In contrast, the path-based method uses the user-object diagram to discover the path similarity between users or objects through predefined meta-paths or automatically mined connection patterns [7]. This method can provide the interpretability of recommendation results. However, the design of reasonable path relies heavily on prior knowledge. At present, the research trend is to combine the embedded method and the path-based method to form a joint method [8]. This method can fully mine two aspects of information and improve the accuracy and personalization of recommendation.

In this paper, a music recommendation model is proposed, which utilizes the knowledge aggregation of important sampling based on GNN to solve the challenges faced by traditional recommendation systems. Our model uses the semantic relations encoded in the knowledge graph to enhance the representation learning process in GNN. By combining domain-specific knowledge and user-item interaction, our model can capture explicit and implicit user preferences, thus leading to more accurate and personalized recommendations. In order to aggregate the neighborhood information related to the current node more accurately, the model uses sampling based on importance to aggregate the neighborhood information of each layer when sampling the neighbors of each entity in KG, thus avoiding the introduction of noise in high-order relationships. In addition, GNN aggregates the representation and neighborhood information of a given entity, which can simulate high-order neighborhood information and capture the potential interests of users. We compared the proposed model with the baselines, and the results prove the effectiveness of our proposed model.

The rest of the paper is organized as follows: the second part provides a review of the related work. Section 3 introduces the proposed recommendation algorithm in detail. Section 4 describes the experimental setup and gives the results. Finally, the section 5 summarizes the full paper and discusses the future research direction.

II. Related Work

Knowledge graph is a heterogeneous graph representation composed of multiple triples, which contains rich implicit information. The existing recommendation based on knowledge graph aims to improve the accuracy of recommendation through knowledge graph. There are three ways to apply knowledge graph : embedding-based method, path-based method and hybrid-based method[5]. Embedding-based recommendation methods usually use knowledge graph embedding method to learn entity representation. CKE [6] used KGE method TransR to extract KG information, and then combined KG representation with text and visual representation to generate multimodal representation for the item. DKN [9] learnd the entities contained in news by KGE method, and embedded these entities into news recommendation to capture the relationship between different news, thus realized the click rate prediction. Path-based method infers user preferences by pre-defining metapaths or automatically mining connection patterns, and KPRN [7] extracted multi-hop inference paths from KG. EIUM [10] captured the semantic path between specific user-item pairs in KG. Then these multi-hop paths were coded and weighted, thus provided path explanation for sequential recommendation. The hybrid-based method combines the embedded-based method and the path-based method. RippleNet [11] collected neighborhood information around users through preference propagation mechanism, obtained users' potential interests. KGCN [12] used GNN to recursively aggregate high-order neighbors to update the node representation, thus modeling the user-item relationship.

III. Methodology

In this section, we will discuss the proposed model KAISG in detail and introduce the related work.

A. Framework

The overall network is as shown in the Figure 2. The input of model KAISG is user-item interaction and knowledge graph , and the output is the predicted click probability of users on candidate items. First, we construct a knowledge graph according to the data set, and the nodes include tracks, genres, albums and other information. These nodes form a heterogeneous knowledge graph . Then we use the TransE [13] method to obtain the representation of the nodes on the knowledge graph , in which $\boldsymbol{e^h}$ represents the embedding of the head entity h, $\boldsymbol{e^{t_1}}$ represents the embedding of the tail entity t_1. Then, the introduction of high-order noise is avoided by importance sampling of the neighbors of the head entity, and then update the node representation by using the aggregated neighbor information iterated by the GNN, so as to obtain the item embedding representation $\boldsymbol{e^h[N]}$, N is the number of iterative aggregation. $\boldsymbol{e^u}$ is the user embedding initialized by the embedding layer. Finally, we use the embedding of users and items to calculate the prediction probability.

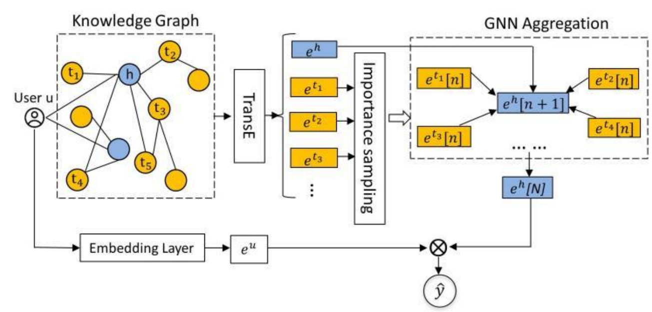

Figure 2. The framework of model KAISG

B. Knowledge Graph Embedding Modeling

TransE is a commonly used knowledge graph embedding method, which is used to map entities and relationships into low-dimensional vector space. It is based on the assumption that the representation of relationships can be obtained by translating the vector representation of entities in the knowledge graph . Specifically, the transition assumes that the relation vector $\boldsymbol{e^r}$ can be obtained by translating the difference vector between the head entity vector $\boldsymbol{e^h}$ and the tail entity vector $\boldsymbol{e^t}$, and the formula is as follows:

$$\boldsymbol{e^h} \approx \boldsymbol{e^r} + \boldsymbol{e^t} \tag{1}$$

The intuition of this assumption is that if the relation r holds, then the difference vector between the head entity h and the tail entity t should be close to the relation vector $\boldsymbol{e^r}$. In order to learn the embedding vectors of entities and relationships, TransE defines an objective function, which optimizes the embedding vectors by minimizing the gap between entities and relationships. Specifically, for a given triplet (h, r, t), the goal of TransE is to minimize the distance between the head entity vector $\boldsymbol{e^h}$ plus the relation vector $\boldsymbol{e^r}$

and the L1 or L2 norm of the tail entity vector $\boldsymbol{e^t}$. The objective function formula is as follows:

$$L=\sum_{(e^h,e^r,e^t)\in T}\sum_{(e^{h'},e^{r'},e^{t'})\in T} max([\gamma+\varepsilon(e^h,e^r,e^t)-\varepsilon(e^{h'},e^{r'},e^{t'})],0) \tag{2}$$

Where $\varepsilon\ (e^h,e^r,e^t)$ is the scoring function, which is defined as follows:

$$\varepsilon\ (e^h,e^r,e^t)=//\ e^h+e^r-e^t\ // \tag{3}$$

By training the objective function, TransE can map entities and relationships into a low-dimensional vector space, in which similar entities and relationships are closer.

C. Importance Sampling

With the expansion of knowledge graph, the number of neighbors of entities will also increase, which may lead to a significant increase in calculation cost and affect the training speed. In addition, the introduction of high-order neighbors may bring some unnecessary noise, thus adversely affecting the recommendation results. In order to solve this problem, our model adopts the method of importance sampling. Specifically, we calculate the importance score of neighbor relationship to users, and select those neighbors with higher value for sampling.

Through importance sampling, we can make more effective use of the information in the knowledge graph , thus improving the accuracy and reliability of the recommendation results. The formula for calculating the importance score of neighbors is as follows:

$$s_r^u=\sigma({e^u}^T e^r) \tag{4}$$

Where $\sigma(x)=\frac{1}{1+exp(-x)}$, $\boldsymbol{e^u}$ is user embedding, $\boldsymbol{e^r}$ is relationship embedding, and r is the relationship between the current entity and its neighbors. After getting the neighbor's importance score. We sample according to the importance scores of neighbors, and sort out the K neighbors with the highest scores for subsequent aggregation.

D. GNN Aggregation

For GNN aggregation, we need to obtain the embedding of the entity itself and its neighborhood embedding. The embedding of the entity itself can be obtained by embedding the knowledge graph , and the neighborhood embedding is the linear combination of the neighbor embedding of the entity. In order to distinguish the importance of different neighbors, when calculating the linear combination of neighbors, we use the normalized importance score as the weight of neighbor embedding. Therefore, the calculation formula of neighborhood embedding is as follows:

$$e^{N(v)}=\sum_{a\in N(v)}\tilde{s}^u_{r_{v,a}}e^a \tag{5}$$

Where $\tilde{s}^u_{r_{v,a}}$ is the normalized importance score, $r_{v,a}$ represents the relationship between entity v and its neighbor a, and $N(v)$ represents the neighbor set of entity v, and the normalization formula is as follows:

$$\tilde{s}^u_{r_{v,a}}=\frac{exp(s^u_{r_{v,a}})}{\sum_{a\in N(v)}exp(s^u_{r_{v,a}})} \tag{6}$$

Finally, we aggregate the entity representation $\boldsymbol{e^v}$ and its neighborhood representation $\boldsymbol{e^{N(v)}}$ into a single vector. We use Sum aggregator, which aggregates by the sum of the embedding of the entity itself and its neighboring embedding, and then carries out nonlinear transformation. The formula of single-layer aggregation is as follows:

$$e^v[1]=\sigma(W\cdot(e^v[0]+e^{N(v)}[0])+b) \tag{7}$$

The above is the case of aggregation of one layer. At this point, $e^v[0]=e^v$, $e^{N(v)}[0]=e^{N(v)}$. $e^v[1]$ is the result of one iteration. Assuming that the number of iterative layers is N, the final aggregation of entity v is expressed as $e^v[N]$, and the formula is as follows:

$$e^v[N]=\sigma(W\cdot(e^v[N-1]+e^{N(v)}[N-1])+b) \tag{8}$$

E. Model Training

The final N -order entity representation is denoted as $\boldsymbol{e^v[N]}$, which is fed into a function $f:R^d\times R^d\rightarrow R$ together with user representation $\boldsymbol{e^u}$ for predicting the probability:

$$\hat{y}=f(e^v[N]^T e^u) \tag{9}$$

We use time back propagation (BPTT) algorithm to train the proposed KAISG model, and the ratio of training set, evaluation set and test set is 6: 2: 2. All parameters can be calculated by gradient descent of back propagation. We choose the Adam optimizer for the momentum.

IV. Experiment

This part is a detailed experimental design, including data set, baselines, evaluation metrics and parameter setting.

A. Data Set

Last.FM contains musicians' listening information from 1872 users of Last.fm online music system. The KG of the data set was built by Microsoft Satori and only keep the entire KG triplet with confidence greater than 0.9. The statistics of the experimental data set is summarized in Table 1.

Table 1. Basic statistics for the data set

Last.FM					
users	items	interactions	entities	relations	triples
1872	3846	21173	9366	60	15518

B. Baselines

- RippleNet[12]: This method makes use of KG information by spreading the user's preference for the entity set on the path in KG with her historical item as the root.
- KGCN[13]: A model combining the characteristics of knowledge graph and graph convolution neural network is proposed, which is based on the multi-hop neighborhood property of the item spreads information to obtain enhanced item representation.

C. Evaluation Metrics

In order to evaluate the performance of our model, we use AUC, ACC and F1 as metrics, and AUC refers to the area under the ROC curve. The ROC curve is a curve formed by TPR as the ordinate and FPR as the abscissa. ACC indicates the ratio of the number of correctly predicted samples to the total number of samples. F1 score is a comprehensive evaluation of precision and recall, and the formula is as follows:

$$F1=\frac{2*recall*precision}{recall+precision} \tag{10}$$

D. Comparison with Baselines

In order to verify the validity of our proposed model, we compare KAISG with the classical baseline. The experimental results of the recommended task are reported in Table 2.

Table 2. Comparison between the proposed method KAISG and baselines

Metrics / Methods	AUC	ACC	F1
Ripplenet	0.7326	0.6870	0.6758
KGCN	0.8558	0.7776	0.7743
KAISG	**0.8598**	**0.7789**	**0.7760**

We can observe that in all the evaluation settings, our KAISG is indeed better than other baselines. Compared with Ripplenet, our model has improved the AUC, ACC and F1 indexes by 17.36%, 13.38% and 14.83% respectively, and compared with KGCN, our model has improved the AUC, ACC and F1 indexes by 0.47%、0.17%、0.22% respectively. This shows that our method can effectively use the knowledge graph to infer the user's preference for music, and at the same time, it shows the effectiveness of our neighbor importance sampling, which can improve the effect of music recommendation.

E. Parameter Setting

In the training stage, we set the embedding dimension of the model as 16 and the batchsize as 128. Negative samples are randomly selected from songs that users have not listened to, and their size is the same as that of positive samples. In addition, the parameters in the model are updated by Adam optimizer, and the learning rate is $5*10^{-4}$.

V. Conclusions

In this paper, a recommendation model KAISG is proposed, which utilizes the knowledge aggregation of important sampling based on GNN to capture the users' preferences. This model enriches the item representation and disseminates high-order information by constructing the topological structure of the item knowledge graph and using GNN to aggregate neighborhood information layer by layer. By means of importance sampling, we can effectively aggregate the neighborhood information related to the current node and avoid the introduction of noise in high-order relationships. Experiments show that our model is superior to the current baseline method on the real data set. In the future work, we hope to verify our model in more KG-based music recommendation scenarios. We also intend to develop different aggregation strategies to integrate the context information of KG to improve the accuracy of recommendation.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No.61773229), the Beijing Municipal Natural Science Foundation (Grant No.KZ201710015010) and Digital Education Research Topic of Beijing (Grant No.BDEC2023619095).

References

[1] SHI Y, LARSON M, HANJALIC A. Collaborative Filtering beyond the User-Item Matrix: A Survey of the State of the Art and Future Challenges[J]. ACM Computing Surveys, 2014: 1-45.

[2] LIU Z, FAN Z, WANG Y, et al. Augmenting Sequential Recommendation with Pseudo-Prior Items via Reversely Pre-training Transformer[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021.

[3] AI Q, AZIZI V, CHEN X, et al. Learning Heterogeneous Knowledge Base Embeddings for Explainable Recommendation[J]. Algorithms, 2018: 137.

[4] HAMILTON WilliamL, YING Z, LESKOVEC J. Inductive Representation Learning on Large Graphs[J]. Neural Information Processing Systems,Neural Information Processing Systems, 2017.

[5] Guo Q, Zhuang F, Qin C, et al. A Survey on Knowledge Graph-Based Recommender Systems[J]. arXiv preprint arXiv:2003.00911, 2020.Jj

[6] ZHANG F, YUAN N J, LIAN D, et al. Collaborative Knowledge Base Embedding for Recommender Systems[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016.

[7] XIAN Y, FU Z, MUTHUKRISHNAN S, et al. Reinforcement Knowledge Graph Reasoning for Explainable Recommendation[C]//Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019.

[8] SHA X, SUN Z, ZHANG J. Attentive Knowledge Graph Embedding for Personalized Recommendation.[J]. Electronic Commerce Research and Applications, 2021.

[9] WANG H, ZHANG F, XIE X, et al. DKN: Deep Knowledge-Aware Network for News Recommendation[J]. Cornell University - arXiv,Cornell University - arXiv, 2018.

[10] HUANG X, FANG Q, QIAN S, et al. Explainable Interaction-driven User Modeling over Knowledge Graph for Sequential Recommendation[C]//Proceedings of the 27th ACM International Conference on Multimedia. 2019.

[11] WANG H, ZHANG F, WANG J, et al. RippleNet: Propagating User Preferences on the Knowledge Graph for Recommender Systems[C]//Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 2018.

[12] WANG H, ZHAO M, XIE X, et al. Knowledge Graph Convolutional Networks for Recommender Systems.[C]//The World Wide Web Conference. 2019.

[13] BORDES A, USUNIER N, GARCIA-DURAN A, et al. Translating Embeddings for Modeling Multi-relational Data[J]. Le Centre pour la Communication Scientifique Directe - HAL - Inria, 2013.