Integrating Social and Knowledge Graphs in GNN-Based Recommender Systems

Zengyi Yu[1,#], Yonggu Wang [1,✉,#], Mengfang Hu[1], Zihan Wang[1]
[1] College of Educational Science and Technology, Zhejiang University of Technology, Hangzhou, China
✉ wangyg@zjut.edu.cn

***Abstract*—Graph Neural Networks (GNN) have emerged as a powerful tool in recommendation systems due to their ability to adeptly model complex relational data. Despite their potential, existing GNN-based approaches often fail to fully harness the synergistic benefits of integrating social networks and knowledge graphs into the recommendation process, overlooking the nuanced differences between these data sources. To address these gaps, we propose a novel Integrating Social and Knowledge Graphs (ISKG) framework tailored for GNN-based Recommender Systems. The ISKG model amalgamates user-item interactions, social connections, and knowledge graph insights into a unified representation, enhancing the recommendation quality through a multi-faceted approach. It starts with generating initial embeddings, progresses through a fusion layer for feature amalgamation, and refines these features in successive propagation layers. An innovative Adaptive Weighting Mechanism dynamically balances the influence of social and knowledge graph-enhanced features, leading to a Prediction Layer that finalizes the recommendations. Our comprehensive evaluation showcases ISKG's superiority over conventional baselines, highlighting its ability to achieve an effective balance between social and knowledge-based recommendations, thus paving the way for more accurate and nuanced recommendation systems. The project details are available at https://yuzengyi.github.io/ISKG/.**

***Index Terms*—Recommendation, Graph Neural Network, Social and Knowledge Graphs**

I. Introduction

In the era of explosive growth of Internet information, recommendation system has successfully evolved into one of the basic tools of information service. It can help users make reasonable choices and decisions, improve the efficiency of data processing, and effectively alleviate the problem of information explosion [1]. The traditional recommendation algorithms are: CF, CB, etc. However, due to the problem of data sparsity, it is impossible to recommend based on more weak correlation information of users. Therefore, scholars have tried to adopt deep learning methods.

In the past few years, deep learning has developed rapidly and has become one of the mainstream methods in artificial intelligence research. In image detection [2], the use of depth perception technology applied to automatic driving [3] and other specific aspects of the application has been rich. The basic idea of deep learning is to obtain a relatively accurate feature representation by stacking multi-layer neural networks and through two steps of linear transformation and nonlinear activation. Among them, graph neural network is widely used in personalized recommendation due to its powerful ability to capture the complex relationship between nodes [4].

The graph neural network (GNN) obtains the representation of each node by iteratively aggregating the information of adjacent nodes of the target node [5]. PATCHY-SAN proposed by Niepert et al [6], LGCN proposed by Gao et al [7], and DCNN proposed by Atwood et al [8]. Before that, transform is also a better way, Transformers excel in handling global context and impacting textures [9]. All use this idea to fully mine the interaction information between node information, improve the sensitivity between graph nodes, and can extract data features in the graph field.

Neural network architectures based on graph structure data such as GCMC [10], GAT [11] and GraphSAGE [12] have achieved remarkable results in some well-known fields (such as social networks and bioinformatics). However, most of the current recommendation methods based on graph structure have two problems : 1) do not consider the social impact of users 2) the correlation between resources is not considered. Social influence refers to the use of interactive behavior data and social relationships between users to build a personalized model that reflects users' social interests, and based on this model to achieve customized content or service recommendations for different users [13]. The knowledge graph contains a large number of entities and their relationships, which can provide richer and more comprehensive information for the recommendation system. By integrating this information into a graph neural network, the model can better understand items, users, and the complex relationships between them.

Based on the user interaction graph, this paper uses the user social relationship graph and the knowledge graph for recommendation. Our goal is to extract high-order semantic information from three graphs to capture the real preferences of each user and the attractiveness of each item. Our main contributions are summarised as follows:

- We developed a new method ISKG, which integrates user item graph (IG), social graph (SG) and knowledge graph (KG) under the framework of graph neural network.
- Extensive experiments on public datasets demonstrate the effectiveness of ISKG and its interpretability in understanding the importance of complex relationships.

✉ Corresponding author
#These authors contributed equally to this work.

The rest of this article is organized as follows. The section II explains the related definitions and model architecture of knowledge graph and social graph. The section III introduces our proposed method in detail. In the section IV, we give the experimental process and results. Finally, we summarize this paper and give the future research directions.

II. Problem Formulation

In the domain of recommender systems, three fundamental objects and three essential types of relationships are recognized.

Initially, we define the two core objects: a user set $U = \{u_1, u_2, \ldots, u_m\}$, where $m = |U|$, and an item set $V = \{v_1, v_2, \ldots, v_n\}$, where $n = |V|$. Each user $u_i \in U$ is represented as a vector that includes attributes such as age, preferences, and other demographic details. Similarly, each item $v_j \in V$ is characterized by a vector that details attributes like genre, price, and other pertinent features.

Subsequently, certain properties of items can be abstracted from the entities and represented as edge information. For instance, a textbook related to mathematics can be represented as (Textbook, BelongsTo, mathematics). Consequently, we introduce an entity set $T = t_1, t_2, \ldots, t_l$, where $l = |T|$.

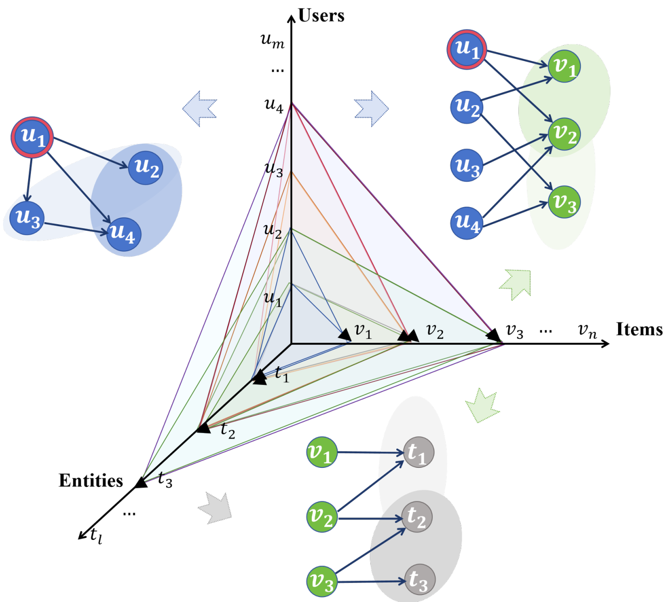

Fig. 1. Illustration of the three-dimensional relationship among users, items, and entities, and its deconstruction into two-dimensional representations: the user-item interaction graph, the social graph, and the knowledge graph.

As elucidated in Fig. 1, we explore the tripartite relationships that form a three-dimensional space and subsequently deconstruct this space into two-dimensional graphs for analysis: the user-item interaction graph, the social graph, and the knowledge graph.

Based on these definitions, we provide a detailed exposition of the problem formulation.

A. User-Item Interaction Graph

Reflecting on user-item relationships, we introduce a user-item interaction matrix $I \in \mathbb{R}^{M \times N}$, which delineates the users' rating preferences for items. Each entry within this matrix signifies a user's level of interest in a particular item. For instance, I_i represents the set of users who have rated item i, with $I_i = \{a | I_{ia} = \text{score}\}$. This notation implies that for item i, the user set I_i consists of users a where the interaction I_{ia} corresponds to a specific rating or score. Consequently, we construct the user-item interaction graph $G_I = (U \cup V, E_I)$, where E_I is the set of edges symbolizing meaningful interactions inferred from the matrix I.

B. Social Graph

Within the context of social relationships, if user u_i follows user u_j, we assign $S_{ij} = 1$; otherwise, $S_{ij} = 0$. This binary schema forms the social matrix S, serving as the user-user adjacency matrix that encapsulates the follow dynamics among users. The social graph is thus represented as $G_S = (U, E_S)$, with E_S emerging from S, illustrating the network of social connections within the user set U.

C. Knowledge Graph

In terms of item-entity relationships, if item v_j is linked to entity t_f, we set $K_{jf} = 1$; otherwise, $K_{jf} = 0$. This binary framework establishes the knowledge matrix K, functioning as the item-entity adjacency matrix. The knowledge graph is therefore denoted as $G_K = (V \cup T, E_K)$, where E_K consists of the set of edges that delineate the relationships between items and entities.

D. Task Description

This section delineates the recommendation task to be undertaken in this study:

- **Input:** a set of graphs (ISKG), encompassing the user-item interaction graph G_I, the social graph G_S, and the knowledge graph G_K.
- **Output:** a predictive model designed to generate a ranked list of items tailored to the user's interests.

III. Methodology

The architecture of the ISKG model, as illustrated in Fig. 2, adopts a hierarchical approach to embedding generation and relevance scoring, with a significant focus on the Propagation Layers (III-C).

The model commences with the Initial Embedding Layer (III-A), which transforms users, items, and entities into compact vector representations, setting the stage for advanced feature integration and refinement.

Central to ISKG is the Fusion Layer (III-B), where a sophisticated amalgamation of features occurs, enhancing the overall representational richness. This layer adeptly merges the inherent and contextual attributes of users and items, fostering embeddings that adeptly reflect the intricate dynamics of user-item interactions.

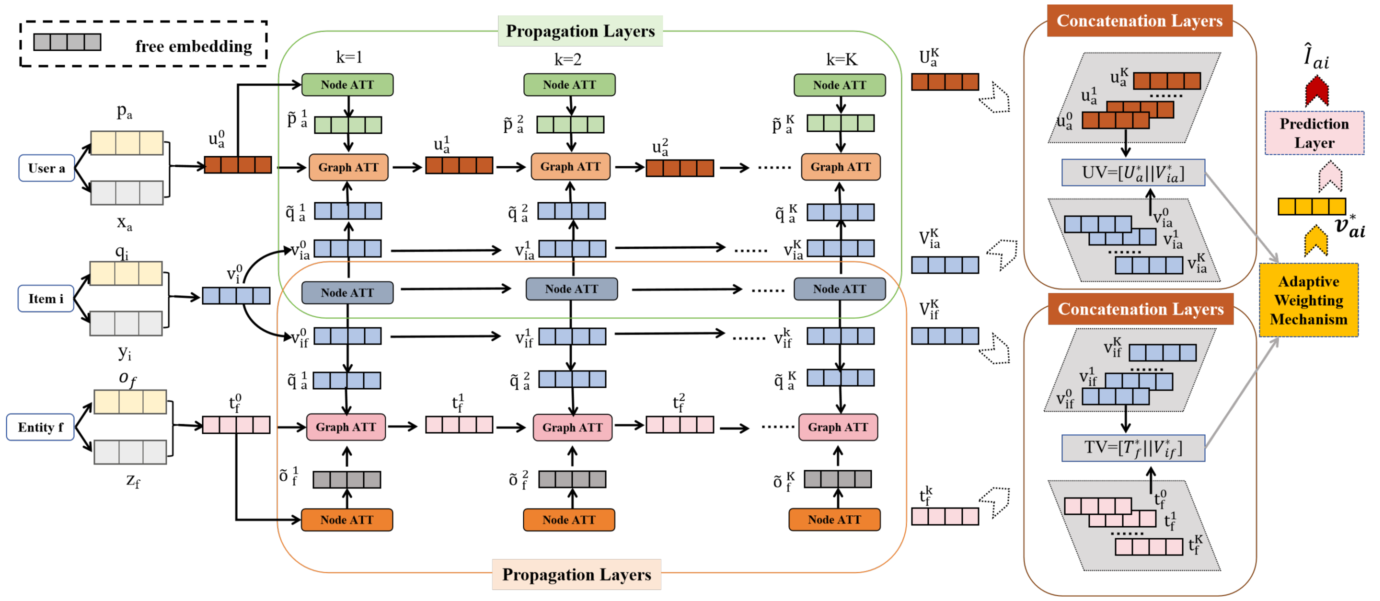

Fig. 2. Architecture of the ISKG Model. Beginning with the derivation of initial embeddings for users, items, and entities, the model progresses through a fusion layer for feature enhancement, followed by iterative refinement in propagation layers. Finally, an Adaptive Weighting Mechanism adjusts the weights of features enhanced by the social graph and knowledge graph, with the Prediction Layer computing the final outcomes.

The crux of ISKG's efficacy, however, lies within the Propagation Layers (III-C). These layers iteratively refine the embeddings, with a deliberate focus on incorporating only the most relevant information at each iteration. This selective assimilation, depicted by the involvement of specific nodes in Fig. 2, amplifies the model's ability to discern and encapsulate the complex web of relationships inherent in the data. By progressively deepening the interaction understanding through these layers, ISKG adeptly captures the nuanced essence of user-item relationships, thereby enhancing the recommendation quality.

A. Initial Embedding Layer

The initial embedding layer is dedicated to encoding user, item, and entity nodes. Its primary function is to transform the input data into low-dimensional embedding vectors $p_u^{(0)} \in \mathbb{R}^d$, $q_i^{(0)} \in \mathbb{R}^d$, and $o_f^{(0)} \in \mathbb{R}^d$ for users, items, and entities, respectively. These embedding vectors are subsequently represented through the embedding matrices P, Q, and O. By integrating the dimensions of P, Q, and O from the ISKG framework, a three-dimensional space can be constructed.

B. Fusion Layer

The Fusion Layer, as the subsequent layer, plays a crucial role in amalgamating features from diverse sources to augment the representational capacity of the model. For each user a, the fusion layer combines the user's embedding vector p_a with its associated feature vector x_a to produce u_a^0:

$$u_a^0 = g(W_a \cdot [p_a, x_a]), \tag{1}$$

where W_a is a transformation matrix, and $g(x)$ is a nonlinear transformation function.

In a similar vein, for each item i, the item embedding v_i^0 is derived as a function of its inherent latent vector q_i and its feature vector y_i, as shown below:

$$v_i^0 = g(W_i \cdot [q_i, y_i]). \tag{2}$$

Mapping v_i^0 onto the P-Q plane yields $v_{ia}^0 = v_i^0$; likewise, when mapped onto the Q-O plane, the result is $v_{if}^0 = v_i^0$.

Finally, for the entity associated with an item, the model construction can be articulated as follows:

$$t_f^0 = g(W_f \cdot [o_f, z_f]). \tag{3}$$

C. Propagation Layers

In each iteration denoted by layer $k+1$, embeddings from users denoted by u_a^k, items denoted by v_i^k and entities denoted by t_f^k from the preceding layer k serve as inputs. The propagation layers successively refine the embeddings to produce v_i^{k+1}, u_a^{k+1}, and t_f^{k+1} through propagation functions. This recursive process initiates at $k = 0$ and culminates upon reaching a predetermined depth K.

Considering item i and its embedding at the kth layer v_i^k, we construct the subsequent embedding v_i^{k+1} at layer $k+1$ from the interest graph G_I as follows:

$$\tilde{v}_{ia}^{k+1} = \text{AGGREGATE}_u(u_a^k,\ \forall a \in I_i) = \sum_{a \in I_i} \eta_{ia}^{k+1} u_a^k, \tag{4}$$

$$v_{ia}^{k+1} = \tilde{v}_{ia}^{k+1} + v_{ia}^k. \tag{5}$$

Here, I_i denotes the set of users who have provided ratings for item i, and u_a^k is the embedding of user a at layer k. The aggregation weight is denoted by n_{ia}^{k+1}, and the updated embedding for each item v_i^{k+1} combines the aggregated neighbor embeddings with the item's previous layer embedding.

This process is analogous for enhancing the embeddings with the knowledge graph.

$$\tilde{v}_{if}^{k+1} = \text{AGGREGATE}_t(t_f^k,\ \forall f \in K_i) = \sum_{f \in K_i} \eta_{if}^{k+1} t_f^k, \quad (6)$$

$$v_{if}^{k+1} = \tilde{v}_{if}^{k+1} + v_{if}^k. \quad (7)$$

Simple mean pooling fails to account for varying significance of user interests in the representation of items. Thus, an attention mechanism is employed to compute the weights η_{ia} and η_{if} in Eqs. (4) and (6), using the following attention functions:

$$\eta_{ia}^{k+1} = \text{MLP}_{\text{attention}_a}([v_{ia}^k, u_a^k]), \quad (8)$$

$$\eta_{if}^{k+1} = \text{MLP}_{\text{attention}_f}([v_{if}^k, t_f^k]). \quad (9)$$

The attention network adopts a MultiLayer Perceptron (MLP) to determine the significance of node connections based on user and item embeddings at layer k. Subsequent normalization of attention weights is performed as follows:

$$\eta_{ia}^{k+1} = \frac{\exp(\eta_{ia}^{k+1})}{\sum_{b \in I_i} \exp(\eta_{ib}^{k+1})}, \quad (10)$$

$$\eta_{if}^{k+1} = \frac{\exp(\eta_{if}^{k+1})}{\sum_{b \in T_f} \exp(\eta_{ib}^{k+1})}. \quad (11)$$

The exponential function ensures the non-negativity of attention weights. Utilizing the exponential function guarantees that each computed attention weight exceeds zero.

For every user a, their latent representation at layer k is symbolized by u_a^k. Central to both the social structure G_S and the affinity graph G_I, a user's embedding u_a^{k+1} at layer $k+1$ assimilates influences from these dual networks: the dispersion of influence within G_S and the dissemination of interests within G_I. The term $\tilde{p}_a^{k+1}$ encapsulates the composite embedding resultant from social adjacency, while $\tilde{q}_a^{k+1}$ aggregates the embedding based on item-centric interest at the subsequent layer. The evolution of each user's embedding u_a^{k+1} is thus articulated as

$$u_a^{k+1} = u_a^k + (\gamma_{a1}^{k+1} \tilde{p}_a^{k+1} + \gamma_{a2}^{k+1} \tilde{q}_a^{k+1}), \quad (12)$$

with $\tilde{p}_a^{k+1}$ and $\tilde{q}_a^{k+1}$ computed through

$$\tilde{p}_a^{k+1} = \sum_{b \in S_a} \alpha_{ab}^{k+1} u_b^k, \quad (13)$$

$$\tilde{q}_a^{k+1} = \sum_{i \in I_a} \beta_{ai}^{k+1} v_i^k, \quad (14)$$

where α_{ab}^{k+1} and β_{ai}^{k+1} denote the scores reflecting social and interest influence, calculated as

$$\alpha_{ab}^{k+1} = \text{MLP}_{\text{social}}([u_a^k, u_b^k]), \quad (15)$$

$$\beta_{ai}^{k+1} = \text{MLP}_{\text{interest}}([u_a^k, v_i^k]). \quad (16)$$

Post-calculation of node-specific attentive weights, these are inputted into the graph attention framework, permitting us to formulate the graph attention weights γ_{al}^{k+1} $(l = 1, 2)$ as

$$\gamma_{a1}^{k+1} = \text{MLP}_{\text{attention}}([u_a^k, \tilde{p}_a^k]), \quad (17)$$

$$\gamma_{a2}^{k+1} = \text{MLP}_{\text{attention}}([u_a^k, \tilde{q}_a^k]). \quad (18)$$

In the enhancement phase involving the knowledge graph, the updated entity embedding t_f^{k+1} at layer $k+1$ is the result of the previous embedding t_f^k augmented with the weighted sum of two components: the refined knowledge representation $\tilde{o}_f^{k+1}$ and the interest-driven entity representation $\tilde{q}_f^{k+1}$, expressed as:

$$t_f^{k+1} = t_f^k + (\gamma_{f1}^{k+1} \tilde{o}_f^{k+1} + \gamma_{f2}^{k+1} \tilde{q}_f^{k+1}), \quad (19)$$

where $\tilde{o}_f^{k+1}$ aggregates the influence from the knowledge aspect, computed by:

$$\tilde{o}_f^{k+1} = \sum_{b \in E_f} \alpha_{fb}^{k+1} t_b^k, \quad (20)$$

and $\tilde{tq}_a^{k+1}$ incorporates the influence from the user interest, determined by:

$$\tilde{tq}_a^{k+1} = \sum_{i \in I_a} \tau_{ai}^{k+1} v_i^k. \quad (21)$$

The attentive weights α_{fb}^{k+1} and τ_{ai}^{k+1} are learned via a dedicated MultiLayer Perceptron for the knowledge graph and user interest, respectively:

$$\alpha_{fb}^{k+1} = \text{MLP}_{\text{knowledge}}([t_f^k, t_b^k]), \quad (22)$$

$$\tau_{ai}^{k+1} = \text{MLP}_{\text{interest}}([u_a^k, v_i^k]). \quad (23)$$

The modulation of the entity embeddings is guided by the attention coefficients γ_{f1}^{k+1} and γ_{f2}^{k+1}, formulated as:

$$\gamma_{f1}^{k+1} = \text{MLP}_{\text{attention}}([t_f^k, \tilde{o}_f^k]), \quad (24)$$

$$\gamma_{f2}^{k+1} = \text{MLP}_{\text{attention}}([t_f^k, \tilde{tq}_a^k]). \quad (25)$$

D. Concatenation Layers

Following the K-level propagation process, we collate the embeddings for users and items, represented by u_a^k and v_i^k across all layers up to $k = \{0, 1, 2, \ldots, K\}$. For each user a, we compile her comprehensive embedding as $u_a^* = [u_a^0 \| u_a^1 \| \ldots \| u_a^K]$, which merges her representations across the layers. Correspondingly, for items in the context of the social graph, we assemble the final embedding $v_{ia}^* = [v_{ia}^0 \| v_{ia}^1 \| \ldots \| v_{ia}^K]$, and similarly, for entities, we construct $t_f^* = [t_f^0 \| t_f^1 \| \ldots \| t_f^K]$ as their cumulative embedding. Likewise, within the framework of the knowledge graph, the conclusive item embedding is formulated as $v_{if}^* = [v_{if}^0 \| v_{if}^1 \| \ldots \| v_{if}^K]$. The Concatenation Layers thus enable the synthesis of feature sets that reflect enhancements from the social (Eq.(26)) and knowledge graphs (Eq.(27)) as follows:

$$UV = [u_a^* \| v_{ia}^*], \tag{26}$$

$$TV = [t_f^* \| v_{if}^*]. \tag{27}$$

E. Adaptive Weighting Mechanism

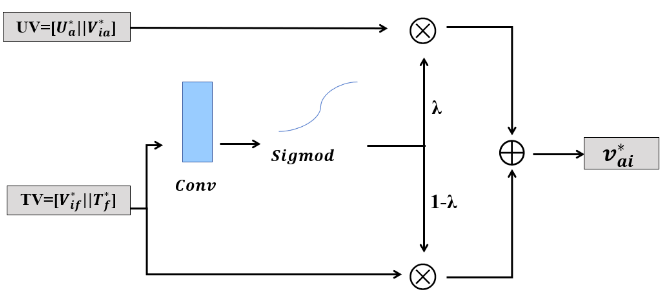

Fig. 3. Schematic of the Adaptive Feature Weighting Module.

Illustrated in Fig.3, The feature vector v_{ai}^* formulates the weighting coefficient λ via a convolutional operation, facilitating the dynamic linkage of UV and TV. The precise formula is articulated as

$$\begin{aligned} \lambda &= W \otimes UV + b, \\ v_{ai}^* &= \varphi(TV) \cdot (1 - \lambda) + UV \cdot \lambda. \end{aligned} \tag{28}$$

where $\varphi(\cdot)$ denotes the process of upsampling, precisely through linear interpolation, and x_i symbolizes the feature map at the ith tier.

F. Prediction Layer

Subsequently, the forecasted score is determined by the dot product of the ultimate user and item embeddings:

$$\hat{I}_{ai} = {v_{ai}^*}^T v_{ai}^*. \tag{29}$$

G. Optimization Method

Our optimization approach utilizes a pairwise ranking loss function, favored in scenarios involving implicit feedback data.

$$L = \min_{\Theta} \sum_{(a,i)\in I^+ \cup (a,j)\in I^-} - \ln(\sigma(\hat{I}_{ai} - \hat{I}_{aj})) + \rho||\Theta||^2, \tag{30}$$

where I^+ signifies the collection of positive instances (known user-item interactions), while I^- represents the negative instances (user-item pairs not observed and hence sampled from the set I). The function $\sigma(x)$ denotes the sigmoid activation, and Θ encompasses the parameters to be optimized. The coefficient ρ is introduced as a regularization term to curb overfitting by penalizing large parameter values. The differentiation property applies to all parameters within this loss function.

The initialization of parameters employs a normal distribution centered at zero with a standard deviation of 0.01. The embedding dimensions within convolutional layers are uniformly maintained, avoiding any intentional alterations. The architecture of the multi-level attention mechanism incorporates a dual-layer MLP. Further details on parameter configurations will be elaborated in the experimental discussions.

IV. EXPERIMENTS

To validate the ISKG framework's robustness, we perform comprehensive experiments on a dataset (Yelp). These experiments are aimed at responding to the following research questions (RQs):

- **RQ1**: How does the ISKG model's performance compare with that of the leading-edge recommendation systems?
- **RQ2**: What are the effects of ISKG's core components on the model's performance?
- **RQ3**: How does altering the weight parameter (λ) impact ISKG's efficacy, especially considering the weight denotes the significance of features enhanced by the social graph and knowledge graph?

A. Experimental Settings

Datasets. To assess the efficacy of the ISKG framework, we conduct evaluations across a dataset. The dataset has been adapted from the KGAT dataset. Detailed statistics for a dataset is presented in Table I.

TABLE I
STATISTICS OF THE DATASETS.

	Yelp
#Users	45,919
#Items	45,538
#Entities	5,538
#User-Item Interaction Graph	1,185,068
#Social Graph	1,035,463
#Knowledge Graph	1,853,704

Utilizing the Yelp dataset, we construct a multi-dimensional graph structure that represents the complex interactions among users, businesses (as items), and their associated attributes (forming entities).

In our **knowledge graph**, businesses are modeled as items, with their location and category attributes extracted to form entities. This extraction process delineates the multifaceted nature of each business and its potential categorization within the graph.

Our **social graph** is derived from the 'friends' attribute within the user data, establishing social connections that reflect the network of interactions and influences among users.

The **user-item interaction graph** is crafted from user reviews, specifically utilizing the star ratings given to businesses. These ratings not only represent user preferences but also serve as the predictive target for our recommendation model's performance evaluation.

Each user, item, and entity is represented as a node within the respective graphs, with edges indicating relationships such as friendships, reviews, or attribute associations. The rating score, which is central to our analysis, is the numerical value that users assign to items, reflecting their satisfaction or experience.

In summary, our dataset amalgamation process for the Yelp dataset entails the following:

- Entities are generated from business attributes, shaping the Knowledge Graph.
- Users are connected through their social ties, forming the Social Graph.
- User-item interactions are established via reviews and star ratings, creating the User-Item Interaction Graph.

This structured approach allows for an extensive analysis of user behavior, social influence, and item characteristics within the domain of recommender systems.

Given the granularity of the Yelp dataset's ratings, we undertake a binary transformation of the score values to facilitate our recommendation model's classification task. Specifically, ratings above the threshold of 3 are converted to a binary value of 1, signifying a positive interaction, while all others are set to 0, indicating a neutral or negative interaction.

Further refining the dataset, we apply a filtering criterion to enhance the quality and reliability of the interactions and social connections represented. Users with fewer than two ratings or social links, as well as items with fewer than two ratings, are excluded from the dataset to ensure a minimum level of engagement and data density.

In the interest of establishing a robust experimental framework, we devise a partitioning scheme that ensures a comprehensive coverage of each user's interaction history. We allocate 80% of each user's recorded interactions to form the training set, which serves as the foundation for the model's learning phase. The residual 20% of interactions are designated as the test set, employed to evaluate the predictive prowess of the model post-training.

Within the training set, we further earmark a random 10% subset of interactions to serve as the validation set. This subset plays a pivotal role in the fine-tuning of hyper-parameters, thus facilitating the optimization of the model's performance prior to its assessment on the test set.

Baselines. In our experiments, we compare the ISKG with various advanced baseline methods:

- **BPR** [14]: A groundbreaking model tailored for item recommendations from implicit feedback, BPR employs a distinct optimization criterion, BPR-Opt, a Bayesian-derived maximum posterior estimator that reshapes the ranking endeavor. Its universally applicable algorithm uses stochastic gradient descent augmented by bootstrap sampling, significantly outstripping conventional learning approaches for personalized ranking endeavors.
- **NFM** [15]: An exemplary factorization model underpinned by neural networks, NFM enhances the FM approach by incorporating an additional hidden layer to enrich the input feature representation, as suggested in [15].
- **CKE** [16]: A method rooted in regularization, CKE leverages semantic embeddings extracted from TransR [17] to fortify the matrix factorization process [14].
- **CFKG** [18]: CFKG employs TransE [19] over a unified graph that integrates users, items, and entities, framing the recommendation system as a prediction of plausible interaction triplets.
- **KGAT** [20]: Building on CFKG, KGAT extracts higher-order semantic connections from the knowledge graph through GNN, allowing for the concurrent refinement of recommendation quality and knowledge graph embeddings.

Evaluation Metrics. In our Top-N recommendation-centric study, the model's efficacy is gauged through two core metrics: Recall@N, and NDCG@N, with N set to 20. Recall@N quantify the correlation, and NDCG@N assesses the ranking accuracy of prevalent items. Enhanced metrics imply superior performance. We conduct each experiment quintuply to ensure equity in evaluation, subsequently averaging the rankings across all items.

Parameter Settings. The model, deployed on PyTorch, operates with a batch size of 1024 and epoch count of 3600. Optimization of parameters is achieved via the Adam optimizer. Optimal learning rates, selected via grid search, are set within $[0.0001, 0.001, 0.005, 0.01]$, and embedding dimensions are chosen from $[16, 32, 64, 128]$. In RQ3, ISKG introduces a manual tuning parameter (λ), pivotal for model adjustment.

Reproducibility. Post-acceptance, our codebase will be hosted on GitHub to foster reproducibility of the study's findings.

B. Recommendation Performance Comparison (RQ1)

In this study, we concentrate on the performance evaluation of the Top-N recommendation system and present the comparative results of our model against other baseline models in Table II.

TABLE II
PERFORMANCE COMPARISON ON YELP DATASET.

	BPR	NFM	CKE	CFKG	KGAT	ISKG	%Improv.
Recall	0.1423	0.1428	0.1437	0.1322	0.1442	**0.1511***	4.79%
NDCG	0.1458	0.1466	0.1405	0.1444	0.1447	**0.1523***	3.89%

- The ISKG model exhibits exemplary performance across all evaluation metrics on the Yelp dataset, underscoring the efficacy of our approach. With respective improvements of 4.79% in Recall and 3.89% in NDCG, the ISKG model validates its superior capability in capturing high-order collaborative signals within user-item interactions and the social domain of users.
- Compared to matrix factorization methods such as BPR and NFM, the ISKG model integrates social relationship information and knowledge graph insights, augmented by an attention mechanism and contrastive learning, to enhance the representational learning of the user social domain. This demonstrates the significance and necessity of GNNs within the recommendation system framework.
- Furthermore, the multi-tier architecture of the ISKG model, as depicted in Fig 2, from the Initial Embedding Layer through to the Fusion and Propagation Layers, provides depth and capacity for the complex user-item relationships. This ensures the model's ability to capture intricate interactions inherent in the data.

C. Ablation Study of ISKG Framework (RQ2)

In order to assess the significance of the core components within the ISKG framework, we conducted an ablation study by creating several model variants:

1) **ISKG-a**: The self-attention mechanism responsible for computing weights in the Propagation Layers is removed.
2) **ISKG-s**: The social enhancement component is omitted.
3) **ISKG-K**: The knowledge graph enhancement component is excluded.

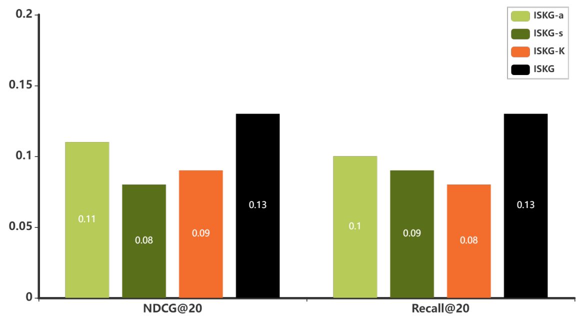

Fig. 4. Impact of Component Ablation on ISKG Performance

The ablation study provides quantitative insights into the importance of individual components within the ISKG framework, as presented in Figure 4.

- The absence of the self-attention mechanism (**ISKG-a**) incurs a relatively minor impact on the model's performance, indicating the robustness of the remaining components. Despite a slight decrement in performance metrics, the model retains a significant portion of its recommendation capabilities.
- A closer examination reveals that the removal of social enhancement (**ISKG-s**) or knowledge graph enhancement (**ISKG-K**) exerts a more pronounced detriment to the model's efficacy. This suggests that these components play a more pivotal role in capturing the nuanced user-item interactions and preferences.
- Interestingly, the comparable influence of omitting social enhancement and knowledge graph enhancement indicates a balanced contribution of these elements to the overall model. The slight variance in their impact underscores the synergistic nature of these enhancements in the ISKG framework.

Based on these observations, it is evident that while the attention mechanism contributes to the fine-grained refinement of user-item interactions, the social and knowledge graph enhancements are integral to the model's holistic understanding of the domain. The nuanced interplay between these components is crucial for the ISKG's superior performance.

D. Impact of Weight Parameter λ on ISKG's Efficacy (RQ3)

In Fig.5, we present the performance trend of ISKG as a function of the weight parameter λ, which balances the influence of features enhanced by the social graph and the knowledge graph.

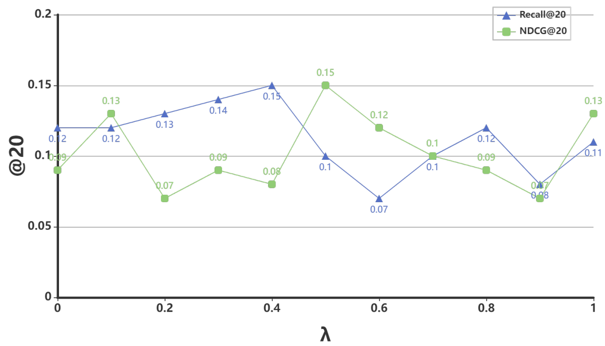

Fig. 5. Variation of ISKG Performance with λ Weight Parameter.

The experimental results, as depicted in Fig.5, lead to several insightful observations:

- The model's performance in terms of NDCG peaks at $\lambda = 0.5$, suggesting an optimal balance between the contributions of the social graph and the knowledge graph at this point. This equilibrium signifies the pivotal role of both social influences and knowledge-based features in the recommendation process.
- Conversely, the highest recall is observed at $\lambda = 0.4$, implying that a slightly greater emphasis on social graph features yields better user-item interaction predictions.
- Notably, at $\lambda = 0.45$, both recall and NDCG metrics converge to similar values, indicating a near-optimal trade-off between the two aspects of the model. This balance might be the key to a more robust and generalizable recommendation system.

The findings underscore the nuanced interplay between different feature enhancements and their collective impact on the recommendation quality. The adaptive weighting mechanism, hence, stands validated as a crucial component of the ISKG framework.

V. Conclusion&Future Work

In this work, we introduce the Integrating Social and Knowledge Graphs in GNN-Based Recommender Systems (ISKG) approach. Specifically, we amalgamate user representations derived from user interest domains, user social domains, and knowledge graphs, thereby facilitating high-order feature propagation characterization. Concurrently, a self-attention mechanism is utilized to refine the learning of salient nodes, thereby enhancing the accuracy in capturing user preferences and item similarities. Notably, we further incorporate an Adaptive Weighting Mechanism to modulate the significance of features augmented by social graphs and knowledge graphs. Extensive experiments and subsequent studies corroborate the efficacy and soundness of the proposed methodology.

Future work will pivot towards exploring noise reduction in the knowledge graph generation process, connecting users and extracted entities to delve into a tripartite space composed of users, items, and entities for a more comprehensive information exploration. Additionally, integrating recommender systems with 3D semantic scene completion [21], could unveil a paradigm where the analysis of users' physical environments enhances personalized recommendations.

Moreover, considering the dynamic changes brought by temporal aspects in user interactions, future versions of our model will aim to incorporate a time decay mechanism. This addition will enable the model to better handle the evolving nature of user preferences over time, potentially offering a more nuanced understanding of temporal dynamics in recommender systems. The time decay mechanism will also serve as a foundational element for enhancing the model's adaptability to changes, ensuring that recommendations remain relevant and timely.

ACKNOWLEDGMENTS

We would like to express our gratitude to the National Natural Science Foundation of China (NSFC) and the Philosophy and Social Science Planning Major Project of Zhejiang Province for their support. This work was supported by the NSFC under Grant No. 62177043 for the project titled "Research on the Emotional and Social Intelligence Perception Model of the Multimodal Feature Fusion Autism Education Robot and Its Applications." Additionally, it was also supported by the Zhejiang Province Philosophy and Social Science Planning Major Project under Grant No. 20XXJC01ZD for the project titled "Construction and Innovation of Smart Services for Social Support of Disabled Persons under 'Internet+'."

REFERENCES

[1] H. Yang, H. Zhou, and Y. Li, "A review of academic recommendation systems based on intelligent recommendation algorithms," in *2022 7th International Conference on Image, Vision and Computing (ICIVC)*, pp. 958–962, IEEE, 2022.

[2] J. Yao, X. Pan, T. Wu, and X. Zhang, "Building lane-level maps from aerial images," *arXiv preprint arXiv:2312.13449*, 2023.

[3] J. Yao and J. Zhang, "Depthssc: Depth-spatial alignment and dynamic voxel resolution for monocular 3d semantic scene completion," *arXiv preprint arXiv:2311.17084*, 2023.

[4] A. Mnih and R. R. Salakhutdinov, "Probabilistic matrix factorization," *Advances in neural information processing systems*, vol. 20, 2007.

[5] Y. Wei, X. Wang, L. Nie, X. He, and T.-S. Chua, "Graph-refined convolutional network for multimedia recommendation with implicit feedback," in *Proceedings of the 28th ACM international conference on multimedia*, pp. 3541–3549, 2020.

[6] M. Niepert, M. Ahmed, and K. Kutzkov, "Learning convolutional neural networks for graphs," in *International conference on machine learning*, pp. 2014–2023, PMLR, 2016.

[7] H. Gao, Z. Wang, and S. Ji, "Large-scale learnable graph convolutional networks," in *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 1416–1424, 2018.

[8] J. Atwood and D. Towsley, "Diffusion-convolutional neural networks," *Advances in neural information processing systems*, vol. 29, 2016.

[9] J. Yao, T. Wu, and X. Zhang, "Improving depth gradient continuity in transformers: A comparative study on monocular depth estimation with cnn," *arXiv preprint arXiv:2308.08333*, 2023.

[10] R. V. D. Berg, T. N. Kipf, and M. Welling, "Graph convolutional matrix completion," 2017.

[11] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

[12] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," 2017.

[13] J. Shokeen and C. Rana, "A study on features of social recommender systems," *Artificial Intelligence Review*, 2019.

[14] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "Bpr: Bayesian personalized ranking from implicit feedback," in *UAI 2009, Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, Montreal, QC, Canada, June 18-21, 2009*, 2009.

[15] X. He and T. S. Chua, "Neural factorization machines for sparse predictive analytics," *ACM SIGIR FORUM*, vol. 51, no. cd, pp. 355–364, 2017.

[16] F. Zhang, N. J. Yuan, D. Lian, X. Xie, and W. Y. Ma, "Collaborative knowledge base embedding for recommender systems," in *the 22nd ACM SIGKDD International Conference*, 2016.

[17] Y. Lin, Z. Liu, M. Sun, Y. Liu, and X. Zhu, "Learning entity and relation embeddings for knowledge graph completion," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 29, 2015.

[18] Q. Ai, V. Azizi, X. Chen, and Y. Zhang, "Learning heterogeneous knowledge base embeddings for explainable recommendation," *Algorithms*, vol. 11, no. 9, 2018.

[19] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," in *Neural Information Processing Systems*, 2013.

[20] X. Wang, X. He, Y. Cao, M. Liu, and T. S. Chua, "Kgat: Knowledge graph attention network for recommendation," *ACM*, 2019.

[21] J. Yao, C. Li, K. Sun, Y. Cai, H. Li, W. Ouyang, and H. Li, "Ndc-scene: Boost monocular 3d semantic scene completion in normalized device coordinates space," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 9455–9465, 2023.