A Method of Social Context Enhanced User Preferences for Conversational Recommender Systems

Zhanchao Gao[1,2], Lingwei Wei[1,2], Wei Zhou[1,2], Meng Lin[1(✉)], and Songlin Hu[1,2]

[1] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
{gaozhanchao, weilingwei, zhouwei, linmeng, husonglin}@iie.ac.cn
[2] School of Cyber Security, University of Chinese Academy of Sciences, Beijing, China

Abstract. Conversational recommender systems (CRS) can dynamically capture user fine-grained preference by directly asking whether a user likes an attribute or not. However, like traditional recommender systems, accurately comprehending users' preferences remains a critical challenge for CRS to make effective conversation policy decisions. While there have been various efforts made to improve the performance of CRS, they have neglected the impact of the users' social context, which has been proved to be valuable in modeling user preferences and enhancing the performance of recommender systems. In this paper, we propose a social-enhanced user preference estimation model (SocialCRS) to leverage the social context of users to better learn user embedding representation. Specifically, we construct a user-item-attribute heterogeneous graph and apply a graph convolution network (GCN) to learn the embeddings of users, items, and attributes. Another GCN is used on the user social context graph to learn the social embedding of users. To estimate better user preference, the attention mechanism is adopted to aggregate the embedding of the user's friends. By aggregating these users' embeddings, we obtain social-enhanced user preferences. Through extensive experiments on two public benchmark datasets in a multi-round conversational recommendation scenario, we demonstrate the effectiveness of our model, which significantly outperforms the state-of-the-art CRS methods.

Keywords: conversational recommender systems · user preference · social context.

1 Introduction

With the advent of intelligent assistants such as Siri (Apple), Alexa (Amazon), and Google Assistant, research on conversational information systems has become increasingly significant.

Conversational recommender systems (CRS) aim to capture dynamic and fine-grained user preferences through interactive conversations with the user [18,

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

19]. In a multi-round conversational recommendation (MCR) scenario, conversational recommendation typically involves two components: recommender component (RC) and conversational component (CC) [17, 28]. The recommender component is responsible for estimating the user's preference for items and attributes, while the conversational component interacts with the user based on the results of the RC and the historical conversational state. Through multi-round conversations with the user, the system can collect rich user feedback, which can clarify the user preference and lead to better recommendation quality.

CRS combines the interactive form of the conversation with recommender systems, thus it can directly ask a user whether he/she likes an item/attribute or not [19] to explicitly obtain the exact preferences of users. Previous studies, such as EAR [17] and SCPR [11] adopt Factorization Machine (FM) [26] as their recommender components to estimate user preference and do not utilize the social context of users. Considering the relation between attribute-level and item-level feedback signals, Xu et al. [32] proposes a user preference estimation model called FPAN in multi-round conversational recommender systems, which does not consider the social context of users either. Among these studies, they only consider the interaction between users and the system and elicit users' current preferences through conversations, ignoring the social context of users, which is known to be helpful for modeling users' potential preferences.

In this study, we aim to address the research gap by investigating the impact of social context on conversational recommender systems. Inspired by social correlation theories such as social influence [23], a user's preference is similar to or influenced by his/her socially connected friends [29]. For example, we always like to share our preferences for movies, music, or book with our friends in reality. Analogous to the fact that users like to spread their preferences with their friends and users' interests are influenced by their friends. Since connected users tend to share similar preferences, we think a user's preferences can not only be estimated from the items he interacted with but also can be inferred from his social context. With the assumption of users' interests are influenced by their friends, we believe that leveraging the social context of users can help CRS better understand their users' preferences and thus provide more accurate recommendations.

In this paper, we focus on the recommender component of conversational recommender systems and propose a novel social-enhanced user preference estimation model (SocialCRS) for multi-round conversational recommendation to better estimate user preferences. Specifically, we leverage the social context of users and user-item interaction history to construct a user-item-attribute heterogeneous graph. And then apply Graph Convolution Network (GCN) [15] on the heterogeneous graph to learn the embeddings of users, items, and attributes. Besides, GCN is also used on the user social context graph to learn the social embedding of users. Then attention mechanism is adopted to aggregate users' social context to estimate better user preference. With the social context information, the CRS can conduct more accurate and personalized recommendations.

In summary, the main contributions of this work are as follows.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

- We propose a social-enhanced user preference estimation model (SocialCRS) to leverage the social context information of users. By extracting the social context sub-graph in the recommender component of conversational recommender systems and learning graph embedding of users, we improved the performance of conversational recommender systems.
- We use the attention mechanism to integrate users' social information to estimate a better representation of user preferences.
- We conduct experiments on two public datasets. Extensive results show that our model significantly outperforms the state-of-the-art methods.

2 Related Work

2.1 Recommender Systems \Conversational Recommender Systems

Traditional recommender systems have achieved much commercial success and are becoming increasingly popular in the era of big data. They often assume a one-shot inter-action paradigm [14], which makes use of historical user-item interactions to estimate user preference on items and work statically. The most representative methods are Matrix Factorization (MF) [16] and Factorization Machine (FM) [26]. With the development of deep neural networks, Neural FM [12] and DeepFM [9] were developed to enhance FM's representation ability by modeling higher-order and non-linear feature interactions. He et al. [12] presented a Neural Collaborative Filtering (NCF) framework by modeling user-item interactions and estimate users' preferences with deep learning methods.

By combing conversational techniques with recommender systems, conversational recommender systems (CRS) were proposed. Therefore, CRS has the natural advantage of obtaining dynamic and fined-grained users' preferences through user online feedback, having become one of the trending research topics for recommender systems. A variety of conversational recommendation task formulations have been proposed.

Interactive recommender systems [11, 30, 2, 34] and critiquing-based recommender systems [25, 3, 21] utilize real-time user feedback on previously recommended items to improve online recommendation strategy, which can be seen as early forms of CRS [7]. Zhang et al. [33] proposed a System Ask–User Respond paradigm for conversation search and recommendation, which made the system can actively ask appropriate questions to understand the user needs. Li et al. [20] developed a conditional generative model of natural language recommendation conversations and make recommendations of movies in the cold-start setting. Christakopoulou et al. [4] presented a large-scale learned interactive recommendation system that asked the users question about the topic and gives item recommendations. CRM [28] integrated recommender systems and dialogue system technologies and used reinforcement learning (RL) to find the policy to interact with the user.

Item and attribute-based conversational recommender systems, such as EAR [17] and SCPR [11], both of them adopted Factorization Machine (FM) [26] as their recommender components to estimate user preference and . Based on a

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

dynamic weighted graph, Deng et al. [5] proposed an adaptive RL framework and model a unified policy learning method to make decisions in CRS. Closely related to our work, Xu et al. [32] proposed a user preference estimation model called FPAN in multi-round conversational recommender systems to capture the relation between users' feedback information.

Same as traditional recommender systems, one of the key tasks of conversational recommender systems is still to correctly understand the users' preferences. Only by correctly understanding the users' preferences, CRS can make better conversational policy decisions and make better recommendations.

2.2 Social Context Information

Extensive research has demonstrated that the user's social context provides additional information that improves the understanding of user behavior in recommender systems and has been widely used in recommender systems. SoRec [22] exploited social context information by decomposing both the user-item rating matrix and the user-social matrix. Jamali et al. [13] proposed the SocialMF model, which introduces a social trust propagation mechanism in the matrix decomposition, making users' preferences indicate users close to their trust. TrustSVD [8] introduced social trust relationships in SVD++ (Singular Value Decomposition), which can be seen as a combination of both SVD++ and SoRec. With the development of graph neural networks, various graph convolution techniques have also been used to model users' social relationships and obtain better representations of user preferences. For example, Fan et al. [6] used graph neural networks and attention mechanisms to capture information between users and users and users and users and items. Wu et al.[31] used graph neural networks to model the social influence propagation process, which enables better representation of users and items.

Although the use of additional information provided by social context has been widely used in traditional recommender systems, the modeling of user preferences in conversational recommender systems does not consider the use of users' social information to enhance user preference representation and thus improve the performance of conversational recommender systems.

3 Preliminaries

3.1 Problem Formulation

In this section, we introduce the notation used to formalize our task. Formally, let $u \in \mathcal{U}$ denotes a user u from the user set $\mathcal{U}$, $i \in \mathcal{I}$ denotes an item i from the item set $\mathcal{I}$, and $a \in \mathcal{A}$ denotes an attribute a from the attribute set $\mathcal{A}$. Each item i is associated with a set of attributes $\mathcal{A}_i \subseteq \mathcal{A}$, which describe its properties.

In an MCR setting, the session starts with a preferred attribute specified by the user. At each turn, the system's recommender component evaluates the user's preference for items and attributes. According to the results of the RC and the

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

historical conversational state, the conversational component decides whether to ask or recommend. When the system chooses to ask the user whether he likes a given attribute, the user replies with binary feedback, either accepting or rejecting the asked attribute. When the system decides to make recommendations to the user, the user gives explicit feedback, indicating whether he likes or dislikes the recommended list of items. The session ends when the user accepts the recommendation or the conversational process reaches the maximum number of turn T.

In this paper, we focus on the recommender component that estimates the users' attribute preference and item preference, which supports the action decision of the conversational component. By utilizing users' social context, the recommender component can better estimate user preference.

4 Methodology

4.1 Model Overview

Fig. 1 illustrates the architecture of the proposed Social-enhanced user preference estimation model for Conversational Recommender Systems (SocialCRS). It consists of an offline representation learning module in which the initial embeddings of all the users, items, and attributes are generated by a graph convolution network. In addition to the online feedback module realized by Xu et al. [32], we have added two additional modules to take advantage of users' social context information. One of the two modules is a graph convolution network on the user social context graph, the other is an attention network to aggregate users' social context. By combining these three modules, we can get social-enhanced user preference representation e_{all}. Finally, user preference on items and attributes are respectively estimated by modeling interactions between social-enhanced user preference representation with the item embedding and attribute embedding.

4.2 Representation Learning

Different from FPAN, SocialCRS not only utilizes the historical user-item interaction data and the relations between items and attributes but also leverages the social connections between users to learn the initial representations of users, items, and attributes. Specifically, a user-item-attribute heterogeneous graph is constructed.

Formally, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denotes the constructed graph, where the set of nodes is denoted as $\mathcal{V} = \mathcal{U} \cup \mathcal{I} \cup \mathcal{A}$, and the set of edges $\mathcal{E}$ consists of four types of edges: the user-item edge (u, i) means the user u interacted with item i, the user-attributes edge (u, a) means the user u prefers the attribute a, and the item-attribute edge (i, a) means that the item i contains the attribute a, and the user-user edge (u, u_f) means the relationship between the user u and his friend u_f.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

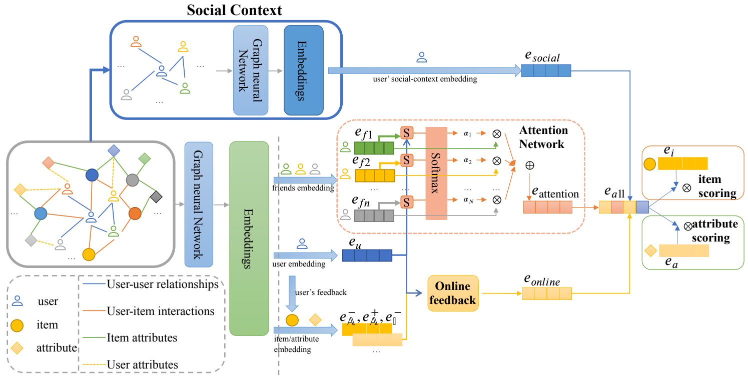

Fig. 1. The overall architecture of the proposed SocialCRS. First, we use GCN on the user-item-attribute heterogeneous graph to pre-train the embeddings of users, items, and attributes. Then, GCN is used on the user social context graph to learn the social embedding of users. The attention mechanism is adopted to aggregate the embedding of users' friends to estimate better user preference. Finally, the final social-enhanced user preference representation is obtained by aggregating three different representations: e_{social}, $e_{attention}$ and e_{online}

Following previous work [32], GraphSAGE [10] is adopted to learn the node representations. For the k_{th} layer representation of node $v \in \mathcal{V}$, GraphSAGE generates node embedding:

$$h_v^k = \sigma(W^k \cdot CONCAT(h_v^{k-1}, h_{N(v)}^k)), \tag{1}$$

where σ means the activate function, W^k is trainable parameters, and h_v^{k-1} is the $k-1_{th}$ layer representation of node v, $N(v)$ denotes the set of node $v's$ neighbors.

Then, to avoid the over-smoothed embedding at the last layer and capture different semantics at different layers, we aggregate the representations generated at different layers and get the final representation of the nodes:

$$e_v = \frac{1}{L+1}\sum_{j=0}^{L} h_v^j, \tag{2}$$

where $v \in \mathcal{V}$. Since $\mathcal{V} = \mathcal{U} \cup \mathcal{I} \cup \mathcal{A}$, we use e_u, e_i, e_a to denote the embeddings of the user, item, and attribute respectively.

4.3 Social-Enhanced User Preference Estimation

Social Context Graph Convolution We have applied GCN on the user-item-attribute heterogeneous graph to learn the node representations. To better learn the user's embedded representation with social context information, we extract the social context sub-graph $\mathcal{G}_{\mathcal{U}}$ from the above heterogeneous graph $\mathcal{G}$. The

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

social context sub-graph $\mathcal{G}_{\mathcal{U}}$ only contains user nodes and edges that consist of social connections between users. GCN can utilize users' social relations naturally and aggregate multi-hop neighbor nodes to learn users' social embedding. Similarly, we adopt GraphSAGE on $\mathcal{G}_{\mathcal{U}}$ to learn users' social embedding as rich context information:

$$e_{social} = \frac{1}{L+1} \sum_{j=0}^{L} h_u^j, \tag{3}$$

where h_u^j denotes the j_{th} layer representation of node $u \in \mathcal{U}$.

Attention Mechanism The attention mechanism has been shown effective in many machine learning tasks. It simulates human recognition by focusing on some selective parts of the whole image or the whole sentence while ignoring some other informative parts [1]. Depending on how close a user is to his or her friends, those friends have different effects on a user's preferences. By regarding the user's friends as an image or a sentence, an attention mechanism is applied to learn the weights of the user's friends' influence on the user's preference. Therefore we can further fuse social influence among users. To be specific, we first calculate the similarity between the user's representation e_u and his friend's representation e_{f_i}:

$$S_i = f(e_u, e_{f_i}) = e_u W^Q \cdot e_{f_i} W^K, \tag{4}$$

where W^Q and W^K are trainable parameters. Then, we use the softmax function to get the weights of the user's friends' influence:

$$\alpha_i = softmax(S_i) = \frac{exp(S_i)}{\sum_i exp(S_i)}. \tag{5}$$

Finally, we get user representation that aggregates the user's social context information:

$$e_{attention} = \textstyle\sum_i \alpha_i e_{f_i}. \tag{6}$$

Online Feedback Inspired by FPAN [32], with the user online feedback, including $\mathcal{A}_u^+, \mathcal{A}_u^-, \mathcal{I}_u^-$, we derive the online user's preference representation by aggregating different kinds of feedback signals:

$$e_{online} = e_u - e_{\mathcal{I}}^- + e_{\mathcal{A}}^+ - e_{\mathcal{A}}^-, \tag{7}$$

where $e_{\mathcal{A}}^+$ denotes the embeddings of the user's positive attribute feedback $\mathcal{A}_u^+$, $e_{\mathcal{A}}^-$ denotes the embeddings of the user's negative attribute feedback $\mathcal{A}_u^-$, and $e_{\mathcal{I}}^-$ denotes the embeddings of the user's negative item feedback $\mathcal{I}_u^-$. More details can be seen from [32].

Social-Enhanced User Preference Now, the final social-enhanced user preference representation is obtained by aggregating above three different representations:

$$e_{all} = \alpha e_{social} + \beta e_{attention} + \gamma e_{online}, \tag{8}$$

where α, β, γ are hyper parameters, and $\alpha + \beta + \gamma = 1$.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

4.4 Item and Attribute Scoring

Since we have gotten the social-enhanced user preference representation e_{all}, SocialCRS next needs to score items and attributes, deciding which items to recommend and which attribute to ask.

Item Scoring Given an arbitrary item $i \in \mathcal{I}$, we predict how likely u will like i in the conversation session by the dot product between item embedding e_i and the aggregated user preference representation e_{all}:

$$f(i, u) = e_{all} \cdot e_i. \tag{9}$$

Attribute Scoring Similarly, given an arbitrary attribute $a \in \mathcal{A}$, the affinity score between the user u and attribute a can be estimated as the dot product between the attribute embedding e_a and e_{all}:

$$g(a, u) = e_{all} \cdot e_a. \tag{10}$$

4.5 Model Training

Since the scoring of items and attributes are independent, we formulate the task of the goal of accurate item scoring and attribute scoring separately. Following previous works [32, 17], the training objective consists of two loss functions: $\mathcal{L}_{item}$ and $\mathcal{L}_{attr}$.

Item Scoring Loss To make accurate item scoring, we optimize the pairwise Bayesian Personalized Ranking (BPR) [27] loss. We use two types of negative samples $\mathcal{D}_1$ and $\mathcal{D}_2$ tailored for MCR:

$$\begin{aligned}\mathcal{L}_{item} = &\sum_{(u,i^+,i^-)\in\mathcal{D}_1} -ln\sigma(f(i^+, u) - f(i^-, u)) + \\ &\sum_{(u,i^+,i^-)\in\mathcal{D}_2} -ln\sigma(f(i^+, u) - f(i^-, u)) + \lambda_\Theta \left\|\Theta\right\|^2,\end{aligned} \tag{11}$$

where

$$\mathcal{D}_1 := \left\{(u, i^+, i^-) | i^- \in \mathcal{I} \setminus \mathcal{I}_u\right\}$$

and

$$\mathcal{D}_2 := \left\{(u, i^+, i^-) | i^- \in \mathcal{I}_{cand} \setminus (\mathcal{I}_u \cup \mathcal{I}_u^-\right\}$$

i^+ denotes the user's target item in a conversation session. $\mathcal{I}_u$ is the set of items historically interacted by user u. $\mathcal{I}_{cand}$ is the candidate item set containing items that satisfy the user's attribute requirements in the current conversation session. And the first loss learns u's general preference, the second loss learns u's specific preference given the current candidates. λ_Θ is the regularization parameter to prevent overfitting.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

Table 1. Dataset Statistics of LastFM and Yelp*

		LastFM	Yelp*
User-Item Interaction	#Users	1801	27675
	#Items	7432	70311
	#Attributes	33	590
	#Interactions	76693	1368606
	#Avg. friend relations per user	13.303	24.868

Attribute Scoring Loss For attribute scoring, we also employ BPR loss, and assume that the attributes of the ground truth item i^+ should be ranked higher than other attributes:

$$\mathcal{L}_{attr} = \sum_{(u,a^+,a^-)\in\mathcal{D}_3} -ln\sigma(g(a^+,u) - g(a^-,u)) + \lambda_\Theta \left\|\Theta\right\|^2, \qquad (12)$$

where the pairwise training data $\mathcal{D}_3$ is defined as:

$$\mathcal{D}_3 := \left\{(u,a^+,a^-)|a^+ \in \mathcal{A}_{i^+} \setminus \mathcal{A}_u^+, a^- \in \mathcal{A} \setminus (\mathcal{A}_{i^+} \cup \mathcal{A}_u^-\right\}.$$

Multi-task training We perform joint training on the two tasks of item scoring and attribute scoring, which has the potential of mutual benefits since their parameters are shared. The multi-task training objective is:

$$\mathcal{L} = \mathcal{L}_{item} + \mathcal{L}_{attr}. \qquad (13)$$

5 Experiments Setups

5.1 Datasets

For better comparison, we follow [17, 19, 32, 5] to conduct experiments on two publicly available datasets: (1) **LastFM**[3] for music artist recommendation and (2) **Yelp**[4] for business recommendation.

Specifically, LastFM contains 1,801 users and 7,432 items, and 76,693 interactions. **Yelp** contains 27,675 users and 70311 items and 1,368,606 interactions. Following the practices in [17], the users that have less than 10 reviews are pruned to reduce the data sparsity. Each of the datasets is split in the ratio of 7:2:1 for training, validation and testing. For the item attributes, Lei et al. [17] preprocess the original attributes of LastFM by manually merging relevant attributes into 33 coarse-grained attributes, and constructing a two-level taxonomy with 29 first-level categories and 590 second-level attributes for Yelp. The Yelp dataset is slightly different from [17, 32] and the same as the dataset used in [19, 5] called **Yelp***, using 590 second-level attributes instead of 29 first-level categories, which can help us reduce heavy manual work and make it more practical. The statistics of the two datasets are presented in Table 1.

[3] https://grouplens.org/datasets/hetrec-2011/
[4] https://www.yelp.com/dataset/

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

5.2 Evaluation Metrics

Following previous work on conversational recommender systems, we use the success rate (SR@T) [28] to evaluate the ratio of successfully recommend the ground truth item by turn T. Besides, we use the average number of turns (AT) to measure the efficiency of conversation. The larger the SR we get, the better recommendation performance we have. And a smaller AT denotes more efficient conversation.

5.3 Baselines

We compare SocialCRS with the following state-of-the-art baselines.

- **Max Entropy** [17]**:** This is a ruled-based method to decide whether to ask or recommend. Each turn it asks the attribute having the maximum entropy among the candidate items or chooses to recommend the top-ranked items with a certain probability. Details can be found at [17].
- **Abs Greedy** [17]**:** This method only has a recommendation component. It only recommends items and updates itself, until it makes a successful recommendation or fails after reaching the maximum number of turns.
- **CRM** [28]**:** This approach is originally designed for the single-round conversational recommendation. It integrates recommender systems and dialogue system technologies and uses reinforcement learning (RL) to find the policy to interact with the user. To adapt this method into the MCR scenario, we follow the description of [17].
- **EAR** [17]**:** The EAR framework consists of three stages, including estimation, action, and reflection, to better converse with users. This method is based on the multi-round conversational recommendation scenario and enhances the interaction between the conversation and recommendation components with a RL framework. FM [26] with attribute-aware BPR [27] is adopted as its recommendation component.
- **SCPR** [19]**:** This method is based on the multi-round conversational recommendation setting, which models conversational recommendation as an interactive path reasoning problem on the graph. It adopts the DQN [24] framework to determine when to ask an attribute or to recommend items, intending to achieve successful recommendations in the fewest turns.
- **UNICORN** [5]**:** This method treats three separated decision-making processes in CRS, including when to ask or recommend, what to ask and which to recommend as a unified policy learning problem. Based on a dynamic weighted graph, UNICORN proposed an adaptive RL framework.
- **FPAN** [32]**:** This work concerns user preference estimation in multi-round conversational recommender systems. It makes use of GNN to learn the offline representations and two gating modules to aggregate the online feedback information, achieving more accurate user preference estimation.

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

5.4 Implementation Details

To maintain a fair comparison, we follow [32] and set the size of the recommendation list as 10, and the maximum turn T as 15. The same reward settings are used in our experiments. The embedding size is set to 64. The number of GraphSAGE layers is set to 2. We optimize the model with Adam optimizer. The L_2 norm regularization is set to be 1e-4. The learning rate is set to 0.005 and 0.001 for the item prediction task and attribute prediction task, respectively.

6 Results and Discussion

6.1 Performance Comparison for Multi-round CRS

Table 2 presents the statistics of the model's performances in terms of SR@15 and AT on two datasets. The results are encouraging. That is, the proposed SocialCRS achieves significantly higher SR and less AT than state-of-the-art baselines, demonstrating the superior performance of SocialCRS. From the reported results, we have the following observations. (1) FPAN and SocialCRS which estimate user preference on both attributes and items, achieve better performance than other baselines. (2) Compared with FPAN, SocialCRS obtains better performance on both datasets. Take the dataset LastFM as an example, SocialCRS gains 9.90% SR@15 improvements and decrease 3.85% AT against FPAN. This validates our social-enhanced user preference estimation in the recommender component helps the conversational component to make better decisions, and improves confidence when making recommendations.

Table 2. Performance comparison of all methods on two datasets by SR@15 and AT, where the best performance is bold-faced. SR@15 is the higher the better, while AT is the lower the better. The number in bold denotes that the improvement of SocialCRS over other methods is statistically significant for $p < 0.05$.

	LastFM		Yelp*	
	SR@15↑	AT↓	SR@15↑	AT↓
Max Entropy	0.283	13.91	0.398	13.42
Abs Greedy [17]	0.222	13.48	0.189	13.43
CRM [28]	0.325	13.75	0.177	13.69
EAR [17]	0.429	12.88	0.182	13.63
SCPR [19]	0.465	12.86	0.489	12.62
UNICORN [5]	0.535	11.82	0.520	11.31
FPAN [32]	<u>0.667</u>	<u>10.14</u>	<u>0.642</u>	<u>10.16</u>
SocialCRS	**0.733(+9.90%)**	**9.75(-3.85%)**	**0.717(+11.68%)**	**9.73(-4.23%)**

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

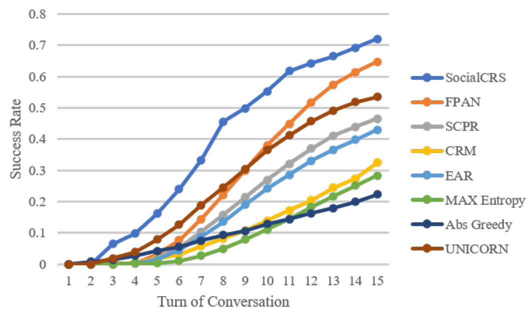

Fig. 2. Success rate of compared methods at different turns on the LastFM dataset.

6.2 Performance Comparison at Different Conversation Turns

In this part, we report the performance comparison of the success rate at each turn (SR@t) in Figure 2.

As can be seen, our SocialCRS model significantly outperforms all baselines in various settings. At almost every turn, SocialCRS has higher success rate. Furthermore, the success rate of SocialCRS starts to promote on the first few turns, earlier than other methods, which indicates the effectiveness of our proposed model.

6.3 Ablation Study

We further conducted an ablation study to show the contributions of different types of integrating the social context of the users on the two datasets. By setting $\alpha = 0$, $\beta = 0$ respectively in (8), we investigate the influence of social context graph convolution (**Social**) and the attention mechanism (**Attention**) on conversational recommender systems.

Table 3 shows the results of the ablation study. From the reported results, we can see that removing any type of integration of the social context of the users results in a performance drop. This indicates that combining both social context graph convolution and the attention mechanism is important to learn user preferences with their social context. Besides, both of them outperform the state-of-the-art baselines, which confirms the effectiveness of introducing social context to estimate user preference for CRS.

6.4 Performance Comparison for User Preference Estimation

In this part, we compare the proposed model with FPAN for user preference estimation of the attribute prediction and item prediction w.r.t AUC score.

Table 4 reports the performance of the attribute prediction and item prediction. From the reported results, we can see that SocialCRS achieves a better performance of preference estimation on both attributes and items, which indicates

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

Table 3. Ablation Study on the two dataset. SR@15 is the higher the better, while AT is the lower the better.

Model	LastFM		Yelp*	
	SR@15↑	AT↓	SR@15↑	AT↓
SocialCRS	**0.733**	**9.75**	**0.717**	**9.73**
- w/o Social	0.710	9.89	0.687	10.04
- w/o Attention	0.716	9.83	0.691	9.90

Table 4. Performance comparison for user preference estimation in terms of the AUC score. The best performance is bold-faced.

Dataset	LastFM		Yelp*	
Preference on	attributes	items	attributes	items
FPAN [32]	0.7852	0.6258	0.9731	0.7771
SocialCRS	**0.7904**	**0.6575**	**0.9837**	**0.7915**

the model learns better user preference with social context information. Specifically, compared to FPAN, on the dataset—LastFM, the AUC of SocialCRS's item prediction increased to 0.6575 while the attribute prediction increased to 0.7904. The same improvement can be observed in the Yelp* dataset. These improvements demonstrate introducing the social context can improve the prediction accuracy not only for item prediction but also for attribute prediction.

7 Conclusion

In this paper, we have investigated the social context to improve the performance of conversational recommender systems. We present a novel social-enhanced model named SocialCRS for integrating user social context to better estimate user preference. Specifically, SocialCRS applies Graph Convolution Network (GCN) to learn the embeddings of users, items, and attributes on a user-item-attributes heterogeneous graph. Additionally, GCN is used on the users' social context graph to learn the social embeddings of users. Futhermore, an attention mechanism is adopted to aggregate users' social context. By aggregating the above three different representations, the proposed model successfully learns better user preferences with the guidance of social context. Extensive experiments show that SocialCRS is a simple yet powerful method to leverage users' social context to estimate better user preferences and provide accurate and personalized recommendations for users in conversational recommender systems.

References

1. Cao, D., He, X., Miao, L., Xiao, G., Chen, H., Xu, J.: Social-Enhanced Attentive Group Recommendation. IEEE Transactions on Knowledge and Data Engineering **33**(3), 1195–1209 (2021)

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

2. Chen, H., Dai, X., Cai, H., Zhang, W., Wang, X., Tang, R., Zhang, Y., Yu, Y.: Large-Scale Interactive Recommendation with Tree-Structured Policy Gradient. In: Proceedings of the AAAI Conference on Artificial Intelligence. pp. 3312–3320 (2019)
3. Chen, L., Pu, P.: Critiquing-Based Recommenders: Survey and Emerging Trends. User Modeling and User-Adapted Interaction **22**(1), 125–150 (2012)
4. Christakopoulou, K., Beutel, A., Li, R., Jain, S., Chi, E.H.: Q&r: A two-stage approach toward interactive recommendation. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 139–148 (2018)
5. Deng, Y., Li, Y., Sun, F., Ding, B., Lam, W.: Unified Conversational Recommendation Policy Learning via Graph-Based Reinforcement Learning. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. p. 1431–1441 (2021)
6. Fan, W., Ma, Y., Li, Q., He, Y., Zhao, E., Tang, J., Yin, D.: Graph neural networks for social recommendation. In: The world wide web conference. pp. 417–426 (2019)
7. Gao, C., Lei, W., He, X., de Rijke, M., Chua, T.S.: Advances and Challenges in Conversational Recommender Systems: A survey. AI Open **2**, 100–126 (2021)
8. Guo, G., Zhang, J., Yorke-Smith, N.: TrustSVD: Collaborative Filtering with Both the Explicit and Implicit Influence of User Trust and of Item Ratings. In: Proceedings of the 29th AAAI Conference on Artificial Intelligence. p. 123–129 (2015)
9. Guo, H., Tang, R., Ye, Y., Li, Z., He, X.: DeepFM: A Factorization-Machine based Neural Network for CTR Prediction. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence. pp. 1725–1731 (2017)
10. Hamilton, W.L., Ying, R., Leskovec, J.: Inductive Representation Learning on Large Graphs. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. p. 1025–1035 (2017)
11. He, C., Parra, D., Verbert, K.: Interactive Recommender Systems: A Survey of the State of the Art and Future Research Challenges and Opportunities. Expert Systems with Applications **56**, 9–27 (2016)
12. He, X., Chua, T.S.: Neural Factorization Machines for Sparse Predictive Analytics. In: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. p. 355–364 (2017)
13. Jamali, M., Ester, M.: A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks. In: Proceedings of the 4th ACM Conference on Recommender Systems. p. 135–142 (2010)
14. Jannach, D., Manzoor, A., Cai, W., Chen, L.: A Survey on Conversational Recommender Systems. ACM Computing Surveys **54**(5), 1–36 (may 2021)
15. Kipf, T.N., Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks. In: 5th International Conference on Learning Representations (2017)
16. Koren, Y., Bell, R., Volinsky, C.: Matrix Factorization Techniques for Recommender Systems. Computer **42**(8), 30–37 (2009)
17. Lei, W., He, X., Miao, Y., Wu, Q., Hong, R., Kan, M.Y., Chua, T.S.: Estimation-Action-Reflection: Towards Deep Interaction Between Conversational and Recommender Systems. In: Proceedings of the 13th International Conference on Web Search and Data Mining. p. 304–312 (2020)
18. Lei, W., He, X., de Rijke, M., Chua, T.S.: Conversational Recommendation: Formulation, Methods, and Evaluation. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. p. 2425–2428 (2020)

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15

19. Lei, W., Zhang, G., He, X., Miao, Y., Wang, X., Chen, L., Chua, T.S.: Interactive Path Reasoning on Graph for Conversational Recommendation. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. p. 2073–2083 (2020)
20. Li, R., Kahou, S., Schulz, H., Michalski, V., Charlin, L., Pal, C.: Towards Deep Conversational Recommendations. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems. p. 9748–9758 (2018)
21. Luo, K., Yang, H., Wu, G., Sanner, S.: Deep Critiquing for VAE-Based Recommender Systems. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 1269–1278 (2020)
22. Ma, H., Yang, H., Lyu, M.R., King, I.: SoRec: Social Recommendation Using Probabilistic Matrix Factorization. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management. p. 931–940 (2008)
23. Marsden, P.V., Friedkin, N.E.: Network Studies of Social Influence. Sociological Methods & Research **22**(1), 127–151 (1993)
24. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al.: Human-level Control Through Deep Reinforcement Learning. Nature **518**(7540), 529–533 (2015)
25. Pu, P., Faltings, B.: Decision tradeoff using example-critiquing and constraint programming. Constraints **9**(4), 289–310 (2004)
26. Rendle, S.: Factorization Machines. In: 2010 IEEE International Conference on Data Mining. pp. 995–1000 (2010)
27. Rendle, S., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L.: BPR: Bayesian Personalized Ranking from Implicit Feedback. In: Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence. p. 452–461 (2009)
28. Sun, Y., Zhang, Y.: Conversational Recommender System. In: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. p. 235–244 (2018)
29. Tang, J., Hu, X., Liu, H.: Social Recommendation: A Review. Social Network Analysis and Mining **3**(4), 1113–1133 (2013)
30. Wang, H., Wu, Q., Wang, H.: Factorization bandits for interactive recommendation. In: Thirty-First AAAI Conference on Artificial Intelligence (2017)
31. Wu, L., Sun, P., Fu, Y., Hong, R., Wang, X., Wang, M.: A neural influence diffusion model for social recommendation. In: Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval. pp. 235–244 (2019)
32. Xu, K., Yang, J., Xu, J., Gao, S., Guo, J., Wen, J.R.: Adapting User Preference to Online Feedback in Multi-Round Conversational Recommendation. In: Proceedings of the 14th ACM International Conference on Web Search and Data Mining. p. 364–372 (2021)
33. Zhang, Y., Chen, X., Ai, Q., Yang, L., Croft, W.B.: Towards Conversational Search and Recommendation: System Ask, User Respond. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management. p. 177–186 (2018)
34. Zhou, S., Dai, X., Chen, H., Zhang, W., Ren, K., Tang, R., He, X., Yu, Y.: Interactive Recommender System via Knowledge Graph-Enhanced Reinforcement Learning. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. p. 179–188 (2020)

ICCS Camera Ready Version 2023
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-36027-5_15