

Attention-Guide Walk Model in Heterogeneous Information Network for Multi-Style Recommendation Explanation

Xin Wang,^{1,3,4} Ying Wang,^{1,2*} Yunzhi Ling^{1,2}

¹Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, China

²College of Computer Science and Technology, Jilin University, Changchun 130012, China

³School of Computer Technology and Engineering, Changchun Institute of Technology, Changchun 130012, China

⁴School of Mathematics and Statistics, Northeast Normal University, Changchun 130024, China

xinwangjlu@gmail.com, wangying2010@jlu.edu.cn, yunzhiling21@163.com

Abstract

Explainable Recommendation aims at not only providing the recommended items to users, but also making users aware why these items are recommended. Too many interactive factors between users and items can be used to interpret the recommendation in a heterogeneous information network. However, these interactive factors are usually massive, implicit and noisy. The existing recommendation explanation approaches only consider the single explanation style, such as aspect-level or review-level. To address these issues, we propose a framework (MSRE) of generating the multi-style recommendation explanation with the attention-guide walk model on affiliation relations and interaction relations in the heterogeneous information network. Inspired by the attention mechanism, we determine the important contexts for recommendation explanation and learn joint representation of multi-style user-item interactions for enhancing recommendation performance. Constructing extensive experiments on three real-world datasets verifies the effectiveness of our framework on both recommendation performance and recommendation explanation.

Introduction

A large number of recommendation models are a black-box which does not provide explanation to the user. However, the ideal recommendation system should focus on explanation, form a virtuous circle through user feedback, and constantly improve the performance. Explainable recommendation system not only unveils the recommendation process, but also helps to improve the effectiveness, persuasiveness and satisfaction of recommendation.

In recent years, the tasks of analyzing and mining in heterogeneous information network (Shi et al. 2017) have attracted more and more attention. The representation learning models integrated with rich semantic information are contributed to improve the recommendation performance in the heterogeneous information network. However, these models are still not good at capturing the interactive factors between users and items and reveal how users make decisions.

Recently, the research on attention mechanism provides an effective way to make a recommendation system more interpretable. The most explainable recommendation methods exploit the various attention network for choosing the explanation components, such as aspect-level (Wang et al. 2018b), review-level (Chen et al. 2018), meta path-level (Hu et al. 2018). Although the above methods have achieved better performance, they still have some inherent limitations. Firstly, those methods based on aspect-level or review-level mostly focused on the interaction of user preference to item features in the review. On the contrary, the seldom works studied on inferring the question what user's attributes drive them to like the item. Secondly, when using the meta path as interactive factor of explanation, since the combination of multiple entities and relationships makes the number of meta paths grow exponentially, random walk sample strategy with obtaining low-quality context is unsuitable for generating recommendation explanation. Thirdly, the single explanation style can not fully explore multiple interactive reasons to infer the implicit interaction between users and items. Therefore, we aim at developing an explainable recommendation with providing different explanation styles.

In this paper, we propose a novel framework, *MSRE*, which leverages attention-guide walk model to explore affiliation relations and interaction relations in heterogeneous information network, and to learn the joint representation of user-item interactions based on multi-style explanation meta paths for recommendation prediction and explanation. The major contributions of this paper are summarized as follows:

- We study the problem of multi-style recommendation explanations, which integrates multiple interactive factors between users and items based on affiliation relation and interaction relation to infer user's decision on the item.
- We exploit the attention-guide walk model in heterogeneous information network to selectively sample discriminative attributes (or features) and representative explanation meta paths for recommendation explanation.
- Extensive experiments on three real-world datasets, namely Movielens, Yelp and Amazon, verify the effectiveness and reasonableness of our recommendation explanation framework.

*Corresponding Author

Copyright © 2020, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Related Work

LFM for Explainable Recommendation. (Zhang et al. 2014) aligned each latent factor with an explicit feature in matrix factorization and tracked the factorization procedure. (Chen, Qin, and Xu 2016) conducted a pair-wise rank learning model based on user-item-feature tensor to predict user preferences on features and items, and provided recommendation explanations. (Wang, Wang, and Yin 2018) built multi-task learning with a joint tensor factorization for explainable recommendation: one task was to model recommendation, another was to model explanation with opinionated content.

Topic Modeling for Explainable Recommendation. (McAuley and Leskovec 2013) proposed the Hidden Factor and Topic model that uses a softmax function that aligns each dimension of item (or user) latent vector to each dimension of topic distribution in LDA. (Wu and Ester 2015) proposed Factorized Latent Aspect model that combines the advantages of collaborative filtering and aspect-based opinion mining. (Ren et al. 2017) proposed a social collaborative viewpoint regression model for predicting item ratings based on user opinion reviews and trusted social relations and the explanations were generated based on the discovered user favorite topics embedded in the viewpoints.

Graph-based Learning for Explainable Recommendation. (He et al. 2015) proposed a tripartite graph model named TriRank to model the user-item-aspect ternary relation and make top-K recommendations with relevant explanations. (Heckel et al. 2017) proposed to conduct co-clustering on the user-item bipartite graph for explanations. The explanations were generated by leveraging user collaborative information, such as user-based and item-based explanations. (Wang et al. 2018b) proposed the tree-enhanced embedding model based on attention network for explainable recommendation, which attempts to combine the generalization of embedding models with the tree-based models.

Knowledge Embedding for Explainable Recommendation. (Ai et al. 2018) used the knowledge graph embedding technique to extract explanations with a novel soft-matching algorithm, where the complicated relationships in the knowledge base are briefly represented by embedding vectors. (Wang et al. 2018a) proposed an end-to-end system to fully utilize the knowledge graph for interpretable recommendations. (Wang et al. 2019) proposed a new method named Knowledge Graph Attention Network, which explicitly models the high-order connectivities in Knowledge graph in an end-to-end fashion.

Problem Statement

In this section, we introduce some related concepts and formalize the problem of recommendation explanation.

Definition 1. Heterogeneous Information Network (Sun et al. 2011) is defined as a graph $G = (V, E)$, each node $v \in V$ has a particular node type $\tau(v) \in \mathcal{O}$, and each edge $e \in E$ has a particular edge type $\phi(e) \in \mathcal{R}$. The types of nodes $|\mathcal{O}| > 1$ and the types of edges $|\mathcal{R}| > 1$.

Definition 2. Meta path (Sun et al. 2011) is defined as a schema $S = (\mathcal{O}, \mathcal{R})$, and is denoted in the form of $o_1 \xrightarrow{r_1}$

$o_2 \xrightarrow{r_2} \dots \xrightarrow{r_{m-1}} o_m$, which defines a composite relation $r = r_1 * r_2 * \dots * r_{m-1}$ between edge types r_1 and r_{m-1} , where $*$ represents a composition operation.

Definition 3. Explanation Style based on Meta Path is a presentation of recommendation explanation based on meta paths with the specific types of nodes and edges. The different types nodes and edges can convey the corresponding semantic in a specific explanation style. The good explanation style can help user understand recommendation results with high readability and credibility. The common explanation styles are classified as four categories, namely collaborative-based, feature-based, social-based and hybrid-based.

As the explanation style based on meta path in MovieLens, we take “U-U-M-M” and “U-M-G-M” as two examples, which involve collaborative-based, social-based and feature-based explanation style. The “U-U-M-M” means that recommended because a similar movie was strongly rated by a user who has friendship with the target user, and the “U-M-G-M” means that recommended because a movie sharing the same genres was similar to the movie that was strongly rated by the target user.

Definition 4. Attention-guided Walk (Lee, Rossi, and Kong 2018) is a walk strategy in information network that attention mechanism is used to steer the walk towards a prioritized neighbor. It can be applied to guide which types of nodes or which nodes should be prioritized to the next step in heterogeneous information network.

PROBLEM. For inputs of our framework, we have the user set $\mathcal{U} = \{u_1, u_2, \dots, u_{|\mathcal{U}|}\}$, the item set $\mathcal{I} = \{i_1, i_2, \dots, i_{|\mathcal{I}|}\}$. The set of user implicit feedback $\mathcal{R} = \{r_1, r_2, \dots, r_{|\mathcal{R}|}\}$, $r_{ij} = 1$ when the interaction is observed, and $r_{ij} = 0$ otherwise. For each user $u \in \mathcal{U}$, it consists of a set of attributes $\mathcal{A}_u = \{a_1, a_2, \dots, a_{|\mathcal{A}|}\}$. Given an item $i \in \mathcal{I}$, it includes a set of features $\mathcal{F}_i = \{f_1, f_2, \dots, f_{|\mathcal{F}|}\}$. We first construct a heterogeneous information network and then sample a series of explanation meta paths between user u and item i as the explanation meta path set $\mathcal{P}_{(u,i)}$.

For outputs of our framework, we provide a top-K list of items based on the predicted probability $\hat{r}_{u,i}$ that user u would like item i and the multi-style recommendation explanations $\mathcal{E}_{(u,i)}$.

Our framework has two main tasks: 1) determine the multiple types of user-item interactions for recommendation explanation; 2) learn the joint representation of user-item interactions to make recommendation prediction.

The Proposed Model

Different from the existing recommendation explanation models, we first construct a heterogeneous information network and explore the distinction of various structural relations on meta paths. Since not all meta paths have a positive effect on inferring user-item implicit interaction and various structural relations have significantly different structural characteristics (Lu et al. 2019), we respectively exploit the proposed attention-guided walk model on affiliation relations and interaction relations to determine which node should be prioritized for next step and pick the critical explanation meta paths. Finally, we explore a united way to

learn the representation of user-item interactions for making recommendation prediction, and take the important explanation meta paths as recommendation explanation. The overall architecture is illustrated in Figure 1.

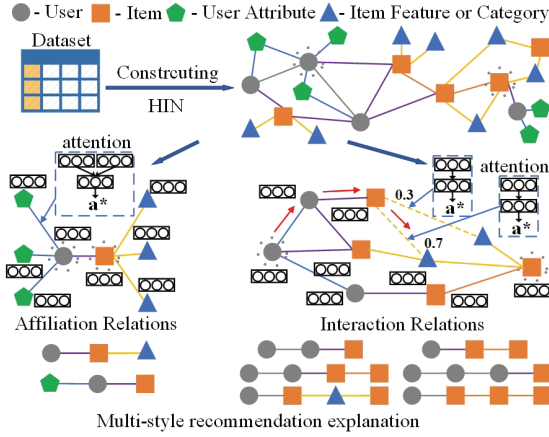


Figure 1: Overall framework of the proposed MSRE

Constructing Heterogeneous Information Network

The purpose of constructing heterogeneous information network is to organize various types of context information for explainable recommendation. Specifically, we extract five types of nodes and organize them into the object set $\mathcal{O} = \{\mathcal{O}_u, \mathcal{O}_i, \mathcal{O}_a, \mathcal{O}_f, \mathcal{O}_t\}$, where $\mathcal{O}_u, \mathcal{O}_i, \mathcal{O}_a, \mathcal{O}_f, \mathcal{O}_t$ respectively denote the sets of users, items, user’s attributes, item’s features and item’s categories. The relationship of these nodes is represented as the relation set $\mathcal{R} = \{\mathcal{R}_{uu}, \mathcal{R}_{ii}, \mathcal{R}_{ui}, \mathcal{R}_{ua}, \mathcal{R}_{if}, \mathcal{R}_{it}\}$, where \mathcal{R}_{uu} denotes user-user similarity or social relation. \mathcal{R}_{ii} denotes item-item similarity relation. \mathcal{R}_{ui} means user-item rating relation. \mathcal{R}_{ua} represents user-attribute association relation. \mathcal{R}_{if} represents item-feature association relation. \mathcal{R}_{it} represents item-category association relation. Due to the limitation of some datasets, we regard item’s features and item’s categories as the same type.

Sampling Explanation Meta Paths

The representation learning based on meta path is one of the most effective way for heterogeneous network embedding. Since only employing the single model for all relations without distinction (Lu et al. 2019), these methods restrict the capability of network embedding. Therefore, when sampling explanation meta paths in heterogeneous information network, we also introduce two different structural relations, which includes affiliation relation (one-centered-by-another structure) and interaction relation (peer-to-peer structure).

For affiliation relation, user-attributes and item-features have typical one-centered-by-another structural characteristics. We respectively take user and item as centered nodes and sample two types of explanation meta paths, namely *user-item-features* and *item-user-attributes*. The length of explanation meta path with affiliation relation is defined as

3. Intuitively, those nodes shared much more similar properties via affiliation relation, are more likely to be the important factors of recommendation explanations.

For interaction relation, since long meta paths are likely to introduce noisy semantics, we only preserve the semantic-specified explanation meta paths, the context nodes are interconnected with both targeting users and recommended items with the maximum length 4. These semantic-specified explanation meta paths are shown in Figure 1. Even then, there are a great number of candidate explanation meta paths. Therefore, we first leverage metapath2vec (Dong, Chawla, and Swami 2017) to pretrain the representation of all nodes in explanation meta paths and replace the nodes with the embedding formation. Then, we define an association degree (similarity) between two nodes in the latent space. Finally, given a type-specific explanation meta path, we take top K ones with the highest average association degree.

Attention-guide Walk on Affiliation Relations

User attributes and item features have the typical characteristic of affiliation relation. Attributes and features as the factors of recommendation explanation is easy to be understood. We explore affiliation relations to preserve the first-order proximity of the target nodes and infer the implicit user-item interactions. We exploit the *Affiliation Relation Attention network* to capture the attention weights on the item’s features and user’s attributes.

Firstly, we project original inputs into the k -dimensional latent representation, namely $\mathbf{u}_i \in \mathbb{R}^k$ for the i th user’s latent vector, $\mathbf{v}_j \in \mathbb{R}^k$ for the j th item’s latent vector, a set of the i th user attribute vectors $\mathbf{a}_i = \{\mathbf{a}_{i1}, \mathbf{a}_{i2}, \dots, \mathbf{a}_{i|A|}\}$ and a set of j th item feature vectors $\mathbf{f}_j = \{\mathbf{f}_{j1}, \mathbf{f}_{j2}, \dots, \mathbf{f}_{j|F|}\}$.

In the affiliation relation attention network, considering a user’s preference is not equally distributed over item’s different features, we formulate the user u_i attention weights α_{ijf} over item i_j and the f th item’s features by the score function $f_{uif}(\mathbf{u}_i, \mathbf{v}_j, \mathbf{f}_f)$ as:

$$\begin{cases} \alpha_{ijf} = \mathbf{h}_f^\top \text{ReLU}(\mathbf{W}_f([\mathbf{u}_i, \mathbf{v}_j \oplus \mathbf{f}_f]) + \mathbf{b}_f) \\ \alpha'_{ijf} = \frac{\exp(\alpha_{ijf})}{\sum_{f' \in \mathcal{F}} \exp(\alpha_{ijf'})} \end{cases} \quad (1)$$

where \mathbf{W}_f and \mathbf{b}_f denote item features’ weight matrix and bias vector, respectively. \mathbf{h}_f is weight vector on feature. \oplus denotes concatenation operation.

Similarly, considering an item’s popularity is also not equally distributed over user’s different demographics, we formulate the item i_j attention weights α_{jia} over user u_i and the a th user’s attributes by score function $f_{jia}(\mathbf{v}_j, \mathbf{u}_i, \mathbf{a}_i)$ as:

$$\begin{cases} \alpha_{jia} = \mathbf{h}_a^\top \text{ReLU}(\mathbf{W}_a([\mathbf{v}_j, \mathbf{u}_i \oplus \mathbf{a}_i]) + \mathbf{b}_a) \\ \alpha'_{jia} = \frac{\exp(\alpha_{jia})}{\sum_{a' \in \mathcal{A}} \exp(\alpha_{jia'})} \end{cases} \quad (2)$$

where \mathbf{W}_a and \mathbf{b}_a denote user attributes’ weight matrix and bias vector, respectively. \mathbf{h}_a is weight vector on attribute. \oplus denotes concatenation operation.

On the basis of learned attention weights, we establish the unified embedding $\mathbf{c}_{u,a}$ and $\mathbf{c}_{i,f}$ by merging user-attributes embedding and item-features embedding, respectively:

$$\mathbf{c}_{u,a} = \sum_{a=1}^{|A|} \alpha_{jia} (\mathbf{u}_i \odot \mathbf{a}_{ia}) \quad (3)$$

$$\mathbf{c}_{i,f} = \sum_{f=1}^{|F|} \alpha_{ijf} (\mathbf{v}_j \odot \mathbf{f}_{jf}) \quad (4)$$

where $\mathbf{c}_{u,a}$ denotes the interaction vector of affiliation relation based on user preference to item driven by the user's attributes, $\mathbf{c}_{i,f}$ denotes the interaction vector of affiliation relation based on the item's features concerned by the user. \odot denotes element-wise product operation.

Attention-guide Walk on Interaction Relations

Different from affiliation relations, interaction relations explore and preserve the high-order proximity between the recommended items and the target users. By sampling meta paths with interaction relations, we can collect the set $\mathcal{P}(u, i)$ of multi-style explanation meta paths. These explanation meta paths can produce a subgraph that starts with the user u and ends with the item i . We leverage the attention-guided walk on interaction relations to determine the further exploring intermediate nodes and choose the critical explanation meta paths.

Intuitively, we regard explanation meta path as a semantic sequence data of user-item interaction. Each intermediate node is the equivalent of an input unit of the sequence data. We take the LSTM model as a core recurrent unit to process new information from the current step and integrate it into new internal representation with information retained from the previous steps. At the current step, we exploit the *Interaction Relation Attention Network* to decide which neighbour to prioritize for further exploration based on the weight of attention distribution. Finally, on the basis of obtaining the attention distributions, we aggregate all hidden state representations of intermediate nodes and generate the embedding of interaction relations.

Specifically, we take an explanation meta path as the input sequence of nodes. The latent representations of these nodes are denoted as $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{|L|}\}$, where $|L|$ is the maximum length of explanation meta path. At the time step t , given the embedding vector of input node \mathbf{x}_t and previous hidden state vector \mathbf{h}_{t-1} , the hidden state of current node \mathbf{h}_t is updated by $\mathbf{h}_t = \text{LSTM}(\mathbf{x}_t, \mathbf{h}_{t-1})$, where the LSTM module is defined as:

$$\begin{aligned} z_t &= \sigma(\mathbf{W}_z \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t]) \\ r_t &= \sigma(\mathbf{W}_r \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t]) \\ \tilde{\mathbf{h}}_t &= \tanh(\mathbf{W} \cdot [r_t * \mathbf{h}_{t-1}, \mathbf{x}_t]) \\ \mathbf{h}_t &= (1 - z_t) * \mathbf{h}_{t-1} + z_t * \tilde{\mathbf{h}}_t \end{aligned} \quad (5)$$

Since the importance of all neighborhood nodes is different, we leverage the *Interaction Relation Attention Network* to determine which neighbors are more important and

should be prioritized for further exploration in the next time step. Specifically, at the time step t , we respectively explore the correlation between the derived hidden state \mathbf{h}_{t-1} and the current nodes \mathbf{x}_t 's neighbor nodes $\{\mathbf{x}_t^1, \mathbf{x}_t^2, \dots, \mathbf{x}_t^{|\mathcal{N}(\mathbf{x}_t)|}\}$, where $\mathbf{x}_t^* \in \mathbb{R}^k$ and $|\mathcal{N}(\mathbf{x}_t)|$ denotes the number of \mathbf{x}_t node's neighbors. Therefore, we define a score function $f_r(\cdot)$ in the form of dot product operation, which measures the closeness between the hidden state \mathbf{h}_{t-1} and the embedding of candidate sampling node \mathbf{x}_t^i ,

$$e_{t,i} = f_r(\mathbf{h}_{t-1}, \mathbf{x}_t^i) \quad (6)$$

where $\mathbf{x}_t^i \in \mathcal{N}(\mathbf{x}_t)$. To reduce the complexity in the learning process, we use the masked attention mechanism to compute $e_{t,i}$ only using the neighborhood of node \mathbf{x}_t .

To make attention weights easily comparable across different neighbors, all weights are normalized for the attention weights $\alpha_{t,i}$ by the softmax function,

$$\alpha_{t,i} = \frac{\exp(e_{t,i})}{\sum_{i'=1}^{|\mathcal{N}(\mathbf{x}_t)|} \exp(e_{t,i'})} \quad (7)$$

In order to show the difference of sampling neighborhood strategy, we can respectively adopt "soft" attention and "hard" attention for determining neighbor nodes and aggregating neighbors context information. The "soft" attention strategy samples all neighbor nodes and aggregates the neighbors with attention distributions. The attention distributions are used to obtain the combination representation from all neighborhood nodes,

$$\mathbf{x}'_t = \sum_{i=1}^{|\mathcal{N}(\mathbf{x}_t)|} \alpha_{t,i} \mathbf{x}_t^i \quad (8)$$

The "hard" attention in (Xu et al. 2015), $s_{t,i}$ is an indicator one-hot variable which is set to 1, the i -th location is used to extract important neighbor. The multinoulli distribution parameterized by $\alpha_{t,i}$ and the final state is defined as,

$$\begin{aligned} p(s_{t,i} = 1 | s_{j < t}, \mathbf{x}_t^i) &= \alpha_{t,i} \\ \mathbf{x}'_t &= \sum_i s_{t,i} \mathbf{x}_t^i \end{aligned} \quad (9)$$

Given a single style explanation meta path p_k , we can obtain the representation of selected neighborhood nodes by the *Interaction Relation Attention Network* and use the sequence $(\mathbf{x}'_{k,1}, \mathbf{x}'_{k,2}, \dots, \mathbf{x}'_{k,|L|})$ as the input of the LSTM model. Then, we adopt the aggregation mechanism to concatenate the embeddings into the final representation,

$$\mathbf{c}_{u,i} = \mathbf{h}_k^{(1)} \parallel \dots \parallel \mathbf{h}_k^{(|L|)} \quad (10)$$

where \parallel is the concatenation operation. $|L|$ denotes the length of explanation meta path.

Joint Model Inference

After performing the attention-guide walk on affiliation relations and interaction relations, we obtain the multi-style recommendation explanation representations $\mathbf{c}_{u \rightarrow i}^S$ between

targeting users and recommended items. Since the single style explanation meta path lacks the expressive power and does not cover all possible user-item interaction factors, we combine multiple recommendation explanation representations based on affiliation relations and interaction relations for improving recommendation performance. Therefore, on the basis of deriving the representations from multi-style explanation meta paths \mathbf{c}_k ($k \in [1, |S|]$, where $|S|$ denotes the number of different style explanation meta paths), we take an *Explanation Style Attention Network* to explore the critical explanation meta paths by introducing the specific representations of user and item. Given the user embedding \mathbf{u}_i , the item embedding \mathbf{v}_j and the embedding of different style explanation meta paths \mathbf{c}_k , the attention weight $\alpha_{i,j,k}$ is calculated by the following the Softmax function,

$$\alpha_{i,j,k} = \frac{\exp(f_p(\mathbf{u}_i, \mathbf{c}_k, \mathbf{v}_j))}{\sum_{k'=1}^{|S|} \exp(f_p(\mathbf{u}_i, \mathbf{c}_{k'}, \mathbf{v}_j))} \quad (11)$$

where $f_p(\cdot)$ is a score function with a simple feed-forward neural network. The attention distributions $\alpha_{i,j,k}$ can be interpreted as the contribution of the different style explanation meta paths to the interaction between user u and item i . After obtaining the attention distributions, the representation of aggregating multi-style explanation meta paths can be given as the following weighted sum,

$$\mathbf{c}_{u \rightarrow i} = \sum_{k=1}^{|S|} \alpha_{i,j,k} \mathbf{c}_k \quad (12)$$

Finally, we combine the user embedding \mathbf{u} , item embedding \mathbf{i} and interaction embedding $\mathbf{c}_{u \rightarrow i}$ into a unified representation and use a MLP network to implement the recommendation prediction,

$$\hat{r}_{u,i} = \text{MLP}(\mathbf{u} \oplus \mathbf{c}_{u \rightarrow i} \oplus \mathbf{v}) \quad (13)$$

Through our framework, we derive multi-style recommendation explanations and enhance their latent representations based on various contexts. The next key problem is how to obtain the final prediction from multi-style recommendation explanations. To answer the problem above, we define an objective function for the joint model inference. Specially, we follow optimization method based on implicit feedback (He et al. 2017) and learn the parameters with negative sampling strategy. The object function of the joint model is defined as,

$$\begin{aligned} L &= - \sum_{(u,i) \in Pos} \log \hat{r}_{u,i} - \sum_{(u,j) \in Neg} \log (1 - \hat{r}_{u,j}) \\ &= - \sum_{(u,i) \in All} r_{u,i} \log \hat{r}_{u,i} + (1 - r_{u,i}) \log (1 - \hat{r}_{u,i}) \end{aligned} \quad (14)$$

where the first term indicates the positive interactions while the second term is the sampled negative interactions.

Experiments

In order to verify the effectiveness of the framework, *MSRE*, we try to answer the following research questions:

Table 1: Statistics of the Three Datasets.

	MovieLens	Yelp	Amazon
# users	943	5117	6427
# items	1682	7213	2716
# features(u/i)	30/18	- /100	- /100
# ratings	100, 000	234, 983	85,607
# user-user	58, 148	168, 580	142, 276
# item-item	80, 586	201, 695	106, 553
# item-category	2, 861	31, 090	18, 587

- Compared with the state-of-the-art methods, how effective is the performance of the proposed framework?
- How much influence does attention-guide walk model on affiliation relations and interaction relations?

Datasets

We adopt three real-world datasets for evaluating our framework, namely MovieLens¹, Yelp² and Amazon³. Due to the limitations of datasets, we take genres as item’s features and item’s categories simultaneously in MovieLens. For Yelp and Amazon, we only consider businesses’ categories and item’ categories as item’s features and retain a fixed number of features for recommended explanation. We assign the higher thresholds to user and item similarities for obtaining the stronger relations in a heterogeneous information network. We summarize the statistics of three datasets in Table 1.

Next, we randomly split the entire ratings into training (70%), validation (10%) and testing (20%) for each dataset. In the training set, for each positive user-item interaction instance, we conduct the negative sampling strategy to pair it with four negative samples that the user did not have any interactions. In the testing set, given one positive instance, we follow the strategy in (He et al. 2017) to reduce time consumption of ranking all items and randomly choose 100 negative samples and pair one positive sample to form a ranked list of the test items.

Experimental Settings

Evaluation Metrics.

We follow the common metrics (Rendle et al. 2009) for top- K recommendation to evaluate the performance, given by *Prec@K*, *Recall@K* and *NDCG@K*. We report the average scores at $K = 10$ of all items in the test set.

Baselines.

We compare the proposed framework with the following representative state-of-the-art methods:

- **ItemKNN** (Sarwar et al. 2001) is a good baseline for implementing item-based collaborative filtering.
- **MF** (Koren, Bell, and Volinsky 2009) is the most representative latent factor model for predicting the rating.

¹<https://grouplens.org/datasets/movielens>

²<https://www.yelp.com/dataset/>

³<https://jmcauley.ucsd.edu/data/amazon/>

- **BPR** (Rendle et al. 2009) is a ranking method based on the pairwise preference assumption for addressing the one-class collaborative filtering problem.
- **NeuMF** (He et al. 2017) is a neural network method by consisting of a generalized MF component and a MLP component for the top- K recommendation.
- **metapath2vec** (Dong, Chawla, and Swami 2017) is a formalizing meta path to construct the embedding of user and item for recommendation. We only consider the patterns of “u-i-u”, “u-i-u-i-u”, “i-u-i” and “i-u-i-u-i”.

Parameter Settings.

All the hyper-parameters are tuned to perform the best on the validation set. We set the embedding dimension $d = 64$. We apply a grid search in $\{0.01, 0.005, 0.001\}$ for learning rate and regularization coefficient 0.0001. To avoid over-fitting, the dropout value is validated from the option set $\{0.1, 0.2, 0.25, 0.5\}$. For sampling meta path, the maximum number of sampled meta path instances is 5.

Experimental Results

Overall Performance.

The comparison results of our framework and the baselines on three datasets are shown in Table 2. After analyzing the results, we make the following conclusions:

(1) The neural network method *NeuMF* achieves better performance than the neighborhood-based collaborate filtering methods (*ItemKNN*, *MF* and *BPR*) (2.1% improvement in precision on average). It suggests that recommendation performance can be improved by using a neural network model. The primary reason is that the deep model can learn the complex correlation between users and items.

(2) The meta path-based method in heterogeneous information network *metapath2vec* outperforms the other baseline methods. It suggests that capturing the proximity context of user and item in a unified way is critical for improving recommendation performance.

(3) In all comparison cases, our *MSRE* model achieves the best performance. On average, it is 2.58% and 1.87% more accurate than *NeuMF* and *metapath2vec*. It indicates that determining important context information (features and explanation meta paths) and exploiting joint representation of multi-style recommendation explanations can improve the performance of recommendation. Moreover, it is also a key factor that attention-guide walk on features and explanation meta paths have a better way of exploring the implicit correlation between user and item.

In summary, we perform a t-test on all comparisons and the t-test results suggest that all improvement is significant. With the help of attention-guide walk model, the proposed framework, *MSRE*, gains significant performance improvement over representative baseline methods, which answers the first question asked at the beginning of this section.

Impact of Attention-guide Walk Model.

In order to study the importance of attention-guide walk model, we construct several variants of our framework on affiliation relations and interaction relations as follows:

- F_{avg} and F_{atten} are two variants based on attention-guide walk on affiliation relations that respectively assign average weight and attention weights to attributes(or features).
- Emp_{avg} , Emp_{hard} and Emp_{soft} are a series of variants based on attention-guide walk on interaction relations that respectively assign average, hard attention and soft attention weights to neighbors. Emp_{rand} randomly samples neighbor, which is the same as random choosing explanation meta path.

As shown in Figures 2 and Figures 3, we have the following observations on Prec@10:

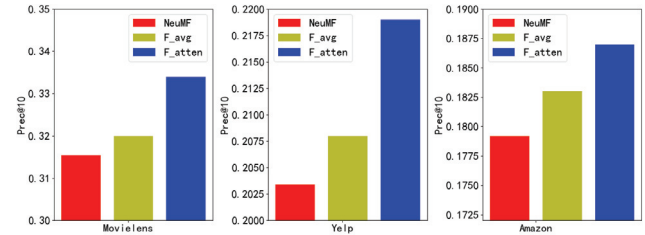


Figure 2: Comparison Results of Multiple Variants on Affiliation Relations

(1) Compare two variants F_{avg} and F_{atten} , since ignoring the importance of different features, F_{avg} without attention mechanism degrades the recommendation performance. It suggests that different features should be assigned different weights according to a specific user-item interaction instead of being treated equally. It verifies that the method of attention-guide walk on features is more effective than only using neural network method *NeuMF*.

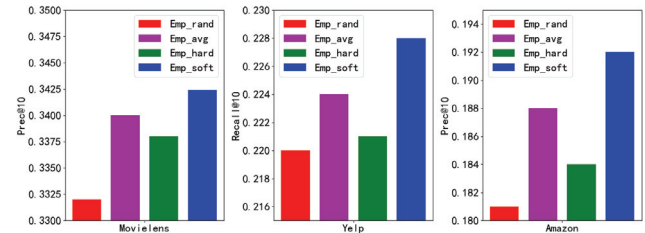


Figure 3: Comparison results of Multiple Variants on Interaction Relations

(2) Emp_{rand} and Emp_{hard} are two variants of only sampling an explanation meta path by randomly and “hard” attention. Emp_{avg} and Emp_{soft} are two variants of aggregating candidate neighbors by mean weights and “soft” attention weights. Obviously, the performance of latter two variants are better than the previous two ones. The main reason is that aggregating the embedding of multiple neighbors fully explores user-item interaction contexts. However, The performance improvement gained by applying attention mechanism into sampling neighbors indicates the advantage of distinguishing different neighbors by attention-guide walk on interaction relations.

Table 2: Performance comparison on three datasets, where the best results are highlighted in bold.

Method	Movielens			Yelp			Amazon		
	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10
ItemKNN	0.2645	0.1619	0.5712	0.1758	0.4761	0.5578	0.1426	0.2845	0.4120
MF	0.3100	0.1952	0.6354	0.1971	0.5014	0.6271	0.1700	0.3183	0.4650
BPR	0.3022	0.1944	0.6235	0.1889	0.4986	0.6013	0.1651	0.3112	0.5146
NeuMF	0.3154	0.2014	0.6495	0.2034	0.5123	0.6348	0.1792	0.3203	0.5217
metapath2vec	0.3213	0.2158	0.6480	0.2168	0.5317	0.6546	0.1811	0.3298	0.5314
MSRE	0.3475	0.2279	0.7049	0.2316	0.5612	0.7120	0.1962	0.3464	0.5458

In summary, the experimental results further demonstrate the importance of attention-guide walk on affiliation relations and interaction relations, which correspondingly answer the second question. In addition, joint representation of context information with the help of attention mechanism can greatly improve the performance of recommendation.

Case Study.

Besides improving performance, the primary motivation of *MSRE* is to generate multi-style recommendation explanation. Therefore, we present some cases in Movielens to highlight the effectiveness of recommendation explanation.

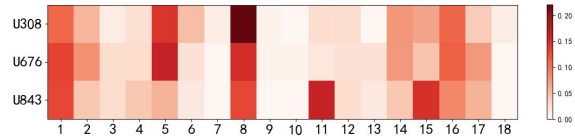


Figure 4: Visualization of three users’ attention weights over genres

Figure 4 shows the heat map of attention weights on features by randomly selecting three users (U308, U676 and U843). We can see that *MSRE* is able to determine the discriminating features. The visual interpretations are that U308 and U676 may prefer the 1th, 5th and 8th genres, while U843 seems to like the 1th, 8th, 11th and 15th genres.

Next, we respectively use the *Interaction Relation Attention Network* and the *Explanation Style Attention Network* to distinguish the neighbors and explanation meta paths. Figure 5 presents the example of attention weights over neighbors and multi-style explanation meta paths.

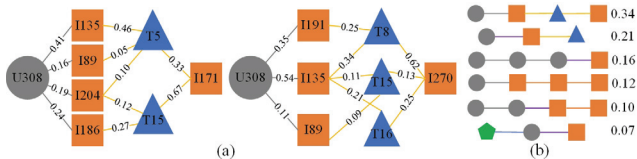


Figure 5: Example of attention weights over neighbors and multi-style explanation meta paths

Figure 5(a) shows two examples of explanation meta path over U308-I171 and U308-I270 with the pattern of “U-I-T-I”. According to attention weights of choosing neighbors,

We can observe that I135 is the most important neighbor for U308, and T5 and T8 are more important neighbors for I135. Figure 5(b) shows the distribution of the attention weights over distinguishing different style recommendation explanations. Take an example of U308-I270, according to the distribution of attention weights, we can observe that the pattern of “U-I-T-I” is more important than the other patterns. It is more suitable as a recommended explanation of the interaction between U308 and I270.

Conclusion

In this paper, we consider the problem of multi-style recommendation explanations based on attention-guided walk model. We propose a novel framework *MSRE* by exploiting recommendation explanations based on the affiliation relations and the interaction relations, and learning their joint representations. The extensive experiments suggest that our framework achieves better performance than other state-of-the-art methods and verifies the effectiveness of recommendation explanation based on attention mechanism.

In future, we will consider to utilize the reinforcement learning to design the attention-guide walk strategy on knowledge graph, and explore the adaptive explanation meta paths and subgraphs for recommendation explanation model in heterogeneous network.

Acknowledgments

This work was supported by a grant from the National Natural Science Foundation of China under grants (No. 61602057, No. 61872161 and No. 61976103), and the China Postdoctoral Science Foundation (No. 2017M611301), and the Jilin Province Key Scientific and Technological Research and Development Project (No. 20190302029GX), and the Nature Science Foundation of Jilin Province (No. 20180101330JC and No. 2018101328JC), and the Foundation of Development and Reform of Jilin Province (No. 2019C053-8), and the Foundation of Jilin Educational Committee (No. JJKH20191257KJ) and the Fundamental Research Funds for the Central Universities, JLU.

References

Ai, Q.; Azizi, V.; Chen, X.; and Zhang, Y. 2018. Learning heterogeneous knowledge base embeddings for explainable recommendation. *Algorithms* 11(9):137.

- Chen, C.; Zhang, M.; Liu, Y.; and Ma, S. 2018. Neural attentional rating regression with review-level explanations. In *Proceedings of the 2018 International Conference on World Wide Web, Lyon, France*, 1583–1592. ACM.
- Chen, X.; Qin, Z.; and Xu, T. 2016. Learning to rank features for recommendation over multiple categories. In *Proceedings of the 39th International Conference on Research & Development in Information Retrieval*, 305–314. ACM.
- Dong, Y.; Chawla, N. V.; and Swami, A. 2017. metapath2vec: Scalable representation learning for heterogeneous networks. In *Proceedings of the 23rd International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada*, 135–144. ACM.
- He, X.; Chen, T.; Kan, M.; and Chen, X. 2015. Trirank: Review-aware explainable recommendation by modeling aspects. In *Proceedings of the 24th International Conference on Information and Knowledge Management, Melbourne, VIC, Australia*, 1661–1670. ACM.
- He, X.; Liao, L.; Zhang, H.; Nie, L.; Hu, X.; and Chua, T. 2017. Neural collaborative filtering. In *Proceedings of the 26th International Conference on World Wide Web, Perth, Australia*, 173–182. ACM.
- Heckel, R.; Vlachos, M.; Parnell, T. P.; and Dünner, C. 2017. Scalable and interpretable product recommendations via overlapping co-clustering. In *Proceedings of the 33rd International Conference on Data Engineering, San Diego, CA, USA*, 1033–1044. IEEE.
- Hu, B.; Shi, C.; Zhao, W. X.; and Yu, P. S. 2018. Leveraging meta-path based context for top-n recommendation with a neural co-attention model. In *Proceedings of the 24th International Conference on Knowledge Discovery & Data Mining, London, UK*, 1531–1540. ACM.
- Koren, Y.; Bell, R.; and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. *Computer* (8):30–37.
- Lee, J. B.; Rossi, R.; and Kong, X. 2018. Graph classification using structural attention. In *Proceedings of the 24th International Conference on Knowledge Discovery & Data Mining, London, UK*, 1666–1674. ACM.
- Lu, Y.; Shi, C.; Hu, L.; and Liu, Z. 2019. Relation structure-aware heterogeneous information network embedding. In *Proceedings of the 33th AAAI Conference on Artificial Intelligence, Honolulu, Hawaii, USA*, 4456–4463. AAAI Press.
- McAuley, J. J., and Leskovec, J. 2013. Hidden factors and hidden topics: understanding rating dimensions with review text. In *Proceedings of the 7th International Conference on Recommender Systems, Hong Kong, China*, 165–172. ACM.
- Ren, Z.; Liang, S.; Li, P.; Wang, S.; and de Rijke, M. 2017. Social collaborative viewpoint regression with explainable recommendations. In *Proceedings of the 10th International Conference on Web Search and Data Mining, Cambridge, United Kingdom*, 485–494. ACM.
- Rendle, S.; Freudenthaler, C.; Gantner, Z.; and Schmidt-Thieme, L. 2009. BPR: bayesian personalized ranking from implicit feedback. In *Proceedings of the 25th International Conference on Uncertainty in Artificial Intelligence, Montreal, QC, Canada*, 452–461. AUAI Press.
- Sarwar, B.; Karypis, G.; Konstan, J.; and Riedl, J. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web, Hong Kong, China*, 285–295. ACM.
- Shi, C.; Li, Y.; Zhang, J.; Sun, Y.; and Philip, S. Y. 2017. A survey of heterogeneous information network analysis. *IEEE Transactions on Knowledge and Data Engineering* 29(1):17–37.
- Sun, Y.; Han, J.; Yan, X.; Yu, P. S.; and Wu, T. 2011. Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. *Proceedings of the Vldb Endowment* 4(11):992–1003.
- Wang, H.; Zhang, F.; Wang, J.; Zhao, M.; Li, W.; Xie, X.; and Guo, M. 2018a. Ripplenet: Propagating user preferences on the knowledge graph for recommender systems. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Torino, Italy*, 417–426. ACM.
- Wang, X.; He, X.; Feng, F.; and Chua, T.-S. 2018b. Tem: Tree-enhanced embedding model for explainable recommendation. In *Proceedings of the 27th International Conference on World Wide Web, Lyon, France*, 1543–1552. ACM.
- Wang, X.; He, X.; Cao, Y.; Liu, M.; and Chua, T. 2019. KGAT: knowledge graph attention network for recommendation. In *Proceedings of the 25th International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA*, 950–958. ACM.
- Wang, N.; Wang, H.; and Yin, Y. 2018. Explainable recommendation via multi-task learning in opinionated text data. In *Proceedings of the 41st International Conference on Research & Development in Information Retrieval, Ann Arbor, MI, USA*, 165–174. ACM.
- Wu, Y., and Ester, M. 2015. FLAME: A probabilistic model combining aspect based opinion mining and collaborative filtering. In *Proceedings of the 8th International Conference on Web Search and Data Mining, Shanghai, China*, 199–208. ACM.
- Xu, K.; Ba, J.; Kiros, R.; Cho, K.; Courville, A. C.; Salakhutdinov, R.; Zemel, R. S.; and Bengio, Y. 2015. Show, attend and tell: Neural image caption generation with visual attention. In *Proceedings of the 32nd International Conference on Machine Learning, Lille, France*, 2048–2057. JMLR.org.
- Zhang, Y.; Lai, G.; Zhang, M.; Zhang, Y.; Liu, Y.; and Ma, S. 2014. Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In *Proceedings of the 37th international conference on Research and Development in Information Retrieval, Gold Coast, QLD, Australia*, 83–92. ACM.