

# Joint User Modeling Across Aligned Heterogeneous Sites Using Neural Networks

Xuezhi Cao<sup>(✉)</sup> and Yong Yu

Apex Data and Knowledge Management Lab,  
Shanghai Jiao Tong University, Shanghai, China  
{cxz, yyu}@apex.sjtu.edu.cn

**Abstract.** The quality of user modeling is crucial for personalized recommender systems. Traditional within-site recommender systems aim at modeling user preferences using only actions within target site, thus suffer from cold-start problem. To alleviate such problem, researchers propose cross-domain models to leverage user actions from other domains within same site. Joint user modeling is later proposed to further integrate user actions from aligned sites for data enrichment. However, there are still limitations in existing works regarding the modeling of heterogeneous actions, the requirement of full alignment and the design of preferences coupling. To tackle these, we propose JUN: a joint user modeling framework using neural network. We take advantage of neural network's capability of capturing different data types and its ability for mining high-level non-linear correlations. Specifically, in addition to site-specific preferences models, we further introduce an auxiliary neural network to transfer knowledge between sites by fine-tuning the user embeddings using alignment information. We adopt JUN for item-based and text-based site to demonstrate its performance. Experimental results indicate that JUN outperforms both within-site and cross-site models. Specifically, JUN achieves relative improvement of 2.96% and 2.37% for item-based and text-based sites (5.77% and 13.54% for cold-start scenarios). Besides performance gain, JUN also achieves great generality and significantly extends the use scenarios.

**Keywords:** Joint user modeling · Personalized recommender system  
Cross-domain recommendation · Neural networks

## 1 Introduction

Personalized recommender system has now become an indispensable tool for the information overload in online environment. A great portion of online services include a native recommender system. Such systems benefit both users and service providers by improving user's online experiences and stimulating potential actions. Driven by its importance, plenty research works emerge in this area. Both research works and online applications prove the success of such systems.

An accurate and comprehensive user modeling is crucial for the quality of personalized recommender systems. Traditional recommender systems capture user preferences by modeling one's previous actions within the target site. For example, recommending movies based on one's movie rating logs. Although the performances are satisfying, problems and limitations still exist.

A well-known and widely exist problem is cold-start problem: recommender systems may fail when dealing with newly arrived users due to lack of historical data [27]. This problem severely jeopardizes user's first impressions. Furthermore, the problem widely exists in almost all recommender systems, making it one of the most urgent problems in this research direction.

Another limitation is the lack of comprehensiveness. When participating in specific online services, user normally reveals only a portion of his preferences. As traditional user modelings focus only on user actions in target site, they can only capture the heavily revealed parts of his preferences. Although these are the most important parts for future recommendations within same site, we cannot claim the other aspects to be totally irrelevant. For example, it is hard to mine one's political preferences using only movie rating histories, but such preferences can help when recommending political movies or documentaries.

To address these problems, researchers propose cross-domain models to leverage user actions in other domains as supplementary data. The intuition behind is that user's underlying preferences is consistent across domains. However, most existing researches focus on domains within same site and only transfer between homogeneous actions (e.g. between different types of movies [28]). There also exist works target at synthetic multi-domain data generated by subdividing single-domain dataset [1, 26]. Although there are works aim at heterogeneous domains, they require aligned actions or additional semantic knowledges [6, 21].

More recently, researchers propose joint user modeling which aims at integrating multiple aligned heterogeneous sites for data enrichment [4, 31]. Given the reality that most people engage in multiple online sites for various needs (Facebook, Twitter, IMDb, etc.), we can integrate user actions in all these sites for a more comprehensive user modeling. Analogous to cross-domain modeling that integrates actions in different domains at site-level, joint user modeling aims at integrating different sites at Internet-level. The advantage is its comprehensiveness and users only face cold-start scenario once when first exposed to the Internet, while the disadvantage is the challenge of heterogeneousness. Due to the task's recency, existing works only focus on fully aligned sites, while a more realistic scenario would be partial alignment or pairwise score-based alignment generated using network aligners [20, 29].

Another common disadvantage of cross-domain recommendation and joint user modeling is that most works employ hard constraints for preferences transfer. Specifically, they follow the framework that first captures user preferences on target and auxiliary domains respectively, and then integrates them using non-personalized hard constraints (e.g. uses user preferences in auxiliary domains as features, or forces the preferences representations to be close). However, these do not match with reality. Although on average users show consistency across

domains (sites), the coupling strength should be identical for each user. For example, there are users who always keep their post the same for Facebook and Twitter, while there also exist users who maintain totally different characters.

In this paper, we employ neural network for joint user modeling to tackle the aforementioned limitations. We favor neural network due to its capability of capturing different types of data, including discrete, sequential, image, etc., which suits the heterogeneous actions in this task. Specifically, we design individual neural networks for each site, capturing the site-specific user preferences. Besides, we design auxiliary neural networks for preferences transfer between sites using the alignment information. All these neural networks share same user embedding layer, thus the auxiliary ones serve as fine-tuning networks using the alignment information. In this paper, we focus on text-based and item-based actions because they are fundamentally heterogeneous and cover most online applications. Therefore, it serves as a good representative for the heterogeneous scenarios. The main contributions of this paper are as follows:

- We propose a neural network based framework JUN for joint user modeling over heterogeneous sites, providing a general solution for this task.
- Within JUN, we propose a novel approach for modeling user preferences consistency across sites using fine-tuning networks, which better transfers knowledge between sites comparing to hard couplings as in existing methods.
- We further employ JUN to model item-based and text-based actions as a representative setting.
- We discuss the integration of JUN with existing online recommender systems to show its utility values besides research contributions.
- Extensive experiments indicate that we achieve a relative improvement of 2.96% and 2.37% for item-based and text-based sites. For cold-start scenarios, we further achieve relative improvements of 5.77% and 13.54%.

The rest of the paper is organized as follows. We first discuss related works in Sect. 2. Then, we present the dataset and conduct preliminary analysis in Sect. 3. We propose JUN in Sect. 4, and report experimental results in Sect. 5. Finally, we summarize and discuss future works in Sect. 6.

## 2 Related Work

### 2.1 User Modeling in Recommender System

Personalized recommender system is now an indispensable tool for finding wanted contents from the overwhelming online data, thus draws plenty research attentions. As user actions vary, researchers propose plenty user modeling techniques accordingly. Matrix factorization [15, 23] is the most widely used method for item-based sites such as e-commercial and movie/music rating sites [13]. Topic models [2] and word embeddings [22] are employed for modeling text-based sites such as Twitter and Tumblr [9]. There are also works targeting at social relationships [18], location-based informations [17], etc.

Recently, several neural network-based recommender systems are proposed. Google proposes Wide & Deep learning which jointly train wide linear models and deep neural networks to combine the benefits of memorization and generalization [5]. He et. al. integrate neural network with matrix factorization, leading to neural collaborative filtering [11]. Neural network's advantage on modeling textual data is also employed for capturing metadata and side-informations [30].

## 2.2 Cold-Start Problem in Recommender System

The cold-start problem is that recommender system may fail when dealing with newly arrived users with no or only few historical actions. Such problem widely exists in almost all recommender systems, hence is of great importance.

Researchers propose various solutions to alleviate this problem [16,27]. One direction is to compensate by additional information such as social relations [8], social tags [32], etc. They require specific types of side information, which are not always available in general. Another direction is to introduce an interview process immediately after registration. This is widely adopted in real applications. Decision trees [10] and functional matrix factorization [34] are employed for generating such questions. Nevertheless, this approach requires additional manual works from users thus still have negative impact on user experiences.

## 2.3 Cross-Domain Recommendation

Following the intuition that user's preferences is consistent across domains, researchers propose cross-domain recommendation to leverage user actions from other domains as auxiliary data [7].

Existing cross-domain approaches mainly focus on homogeneous data. For example, transferring between different types of movies [28]. There is also a good fraction of papers target at synthetic multi-domain data generated by subdividing a single-domain dataset [1,26]. For those scenarios, different 'domains' actually share a lot in common (action type, behavior pattern, etc.).

There do exist works aiming at heterogeneous data. McAuley et al. aim at understanding product ratings with review text [21]. Bayesian hierarchical approach based on LDA is also proposed [28]. Fernandez et. al. further employ semantic information to link items across domains [6]. However, these approaches require aligned actions, or additional semantic informations to help the knowledge transfer, which are not always available in general settings.

## 2.4 Joint User Modeling

Nowadays, as Internet infiltrates into various aspects of our daily life, people always participate in multiple online sites. Following similar intuition of cross-domain recommendation, we can further align multiple sites and integrate the user actions for joint modeling. It extents the site-level integration in cross-domain modeling to an Internet-level integration. Existing works indicate that it is a promising direction for data enrichment [4,31].

Joint user modeling requires the sites to be aligned. Although a large portion of online applications now enable users to login with cross-site accounts (Facebook, Twitter, Google+, etc.), there still exist plenty unaligned users and isolated networks. There are works aiming at automatically alignment by mining personal identifiable information [19], social tags [12], user behaviors [20], etc. An accuracy of over 80% is achieved. Due to the recency of joint user modeling, existing approaches focus only on perfectly aligned sites (with bijection between the user sets). However, when applied to realistic scenarios, we are facing partial alignment or pairwise-score-based alignment recovered by network aligners.

### 3 Data and Preliminary Analysis

#### 3.1 Data Set

The dataset we use is from previous joint user modeling work [4], collected from Douban<sup>1</sup> and Weibo<sup>2</sup>. We exclude users with no actions in one site as well as users marked as inactive, leaving 27,814 users as the final dataset. For item-based site Douban, on average each user has 171.96 movie rating logs. For text-based site Weibo, we have 623.38 microblog actions per user on average, including both original and re-tweets.

#### 3.2 Preliminary Analysis

**User Consistency.** The underlying assumption of joint user modeling is that user preferences are to some extent consistent across sites. As user preferences are not explicitly revealed, we can not measure its consistency directly. Because user actions serve as explicit indicators of user preferences, we give evidence that user action similarity is consistent instead.

Specifically, we analyze whether users with similar actions in one site are still tend to be similar in the other. For each pair of users, we evaluate their similarity according to actions in each site. For text-based site, we first employ Latent Dirichlet Allocation (LDA [2]) to model their topic distributions and then apply KL divergence for similarity. For item-based site, we use Jaccard similarity coefficient upon the sets of rated movies. We show the correlation in Fig. 1(a), where the users are grouped by similarities in source site. Results indicate that user pairs with similar actions in source site also tend to be similar in target site for both directions, which support the assumption.

**Action Count Distribution.** Now we show that leveraging cross-site actions can actually help to enrich the data for cold-start users. Specifically, we analyze whether user’s action count in different sites are highly correlated. If the answer is yes, then cold users would still be cold in parallel sites hence joint user modeling may not lead to a valid data enrichment.

<sup>1</sup> <http://www.douban.com/>.

<sup>2</sup> <http://www.weibo.com/>.

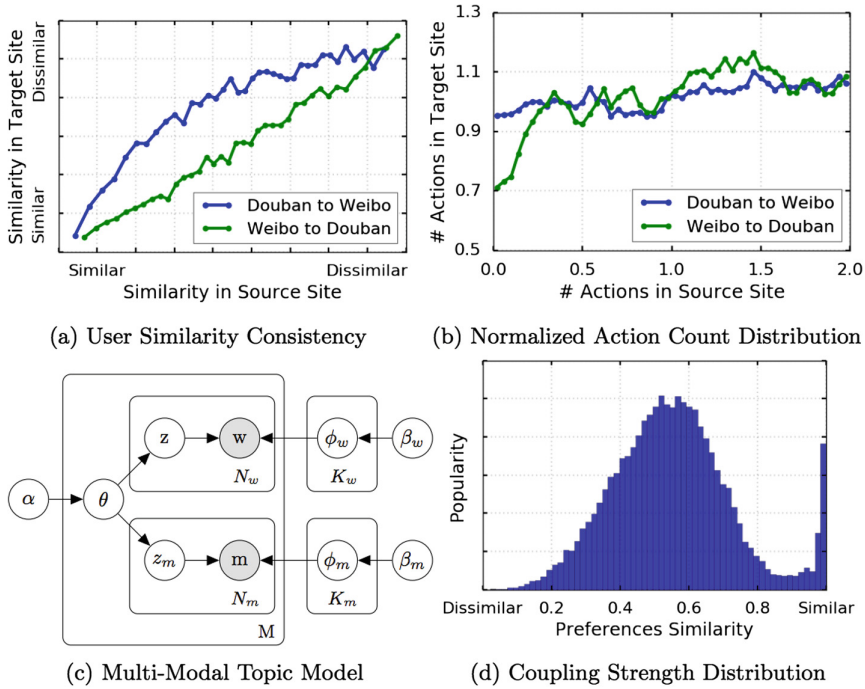


Fig. 1. Preliminary analysis using Weibo-Douban dataset

We group users by normalized action count in source site and report their average normalized action count in target site. Here the normalization is conducted by dividing action count by average action count in corresponding site. We plot the results in Fig. 1(b). Surprisingly, results indicate that user’s action counts in different sites are rather independent. For numeric analysis, the Pearson correlation coefficient is only 0.0549. Therefore, the data enrichment is valid.

**Coupling Strength Distribution.** As stated previously, non-personalized hard coupling methods used in existing works do not match with reality. Now we give evidence by estimating user’s coupling strength distribution between sites.

We first train a multi-modal topic model to represent user preferences in both site using common topic space. We consider movies as another set of ‘words’ and plug them into traditional topic model (showed in Fig. 1(c)). Then, for each user we generate two topic distributions using only microblog actions and movie rating logs respectively. Finally, we estimate coupling strength by cosine similarity between these two topic distributions. We show the distribution in Fig. 1(d). Although the overall trend is towards high similarity, the popularity spreads out among different levels of coupling strength. Therefore, non-personalized hard constraints used in traditional work do not well-represent the reality.

## 4 Our Approach

Our framework JUN is consist of two parts: the models for site-specific preferences and the model for cross-site preferences consistency. Specifically, we have one individual model for each site to capture the user preferences revealed in it. The design of these models depend on site-specific action types and settings, details are given in Sect. 4.1. The cross-site preferences consistency model is designed to transfer knowledge between site-specific models on user representation level using the alignment information, discussed in Sect. 4.2.

User representation is the foundation of user modeling. Most user modeling techniques fall into the framework that model user actions based on user representation  $U_i$  and site-specific information  $A$  (item representation, word embedding, etc.). Formally, these models can be formulated as  $\hat{y}_{ic} = f(U_i, c, A|\Theta)$  where  $\hat{y}_{ic}$  is the estimated action for user  $i$  towards content  $c$ . Take matrix factorization as an example, latent vectors are used to represent both users and items. The prediction can be modeled as  $\hat{y}_{ui} = U_i \cdot A_c$  for the basic setting. User representation provides a good abstraction for user actions. It reveals the required user preferences information while hiding the heterogeneous action details. Therefore, we conduct the integration at user representation level.

In our approach, we employ neural networks for both site-specific preferences model as well as preferences consistency model. Substantial research works successfully employed neural network for varies tasks, including image processing, word embedding, recommender systems, etc. Evidences indicate that neural network has great capability in handling heterogeneous data. Thus, it provides great opportunity for modeling heterogeneous actions. Besides, its ability to capture high-level non-linear correlations is also preferred for preferences consistency modeling. Therefore, we adapt neural network as the underlying technique.

### 4.1 Modeling the Site-Specific Preferences

There exist plenty of heterogeneous action types in different online services. Item-based and text-based actions are typically considered as the most popular ones, covering a wide range of online applications including e-commerce, rating sites, news, blogging, etc. In this work, we focus on modeling movie rating site (item-based) and microblogging site (text-based) as a representative setting.

To integrate site  $\mathcal{S}$  into JUN, we need to define the followings: the user representation  $U_i^{\mathcal{S}}$ , the site-specific information  $A^{\mathcal{S}}$ , and the function  $f^{\mathcal{S}}$  parametrized by  $\Theta^{\mathcal{S}}$  for modeling user actions:  $\hat{y}_{ic}^{\mathcal{S}} = f^{\mathcal{S}}(U_i^{\mathcal{S}}, c, A^{\mathcal{S}}|\Theta^{\mathcal{S}})$ . For generality purpose, we focus on modeling implicit feedback actions, i.e. using only the action itself with no additional information such as rating scores, like/dislike tags, etc. Hence,  $\hat{y}_{ic}^{\mathcal{S}}$  estimates the probability of user  $i$  interacts with content  $c$ .

**Item-Based Site ( $\mathcal{I}$ ):** In item-based sites, actions can be represented as a matrix  $Y \in [0, 1]^{n \times m}$ , where  $n, m$  are the number of users and items respectively,  $Y_{ij}$  indicates whether there is an action from user  $i$  towards item  $j$ .

The most widely used technique for item-based sites is matrix factorization, which models the interaction between user and item by inner product of the

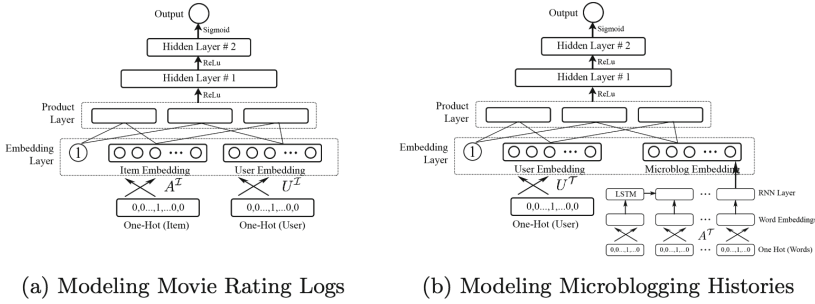


Fig. 2. Neural networks for modeling site-specific preferences

user’s latent vector and the item’s. For probability estimation instead of rating scores, sigmoid function  $\sigma(x) = 1/(1 + e^{-x})$  is often used as the activation function. Formally, we have  $f_{mf}^{\mathcal{I}}(U_i^{\mathcal{I}}, c, A^{\mathcal{I}}) = \sigma(U_i^{\mathcal{I}} \cdot A_c^{\mathcal{I}})$ , where  $U^{\mathcal{I}} \in \mathbb{R}^{n \times k}$ ,  $A^{\mathcal{I}} \in \mathbb{R}^{m \times k}$  and  $k$  is the latent vector dimension.

The underlying methodology of matrix factorization that represents users and items using latent vectors actually matches with embedding layer in neural networks. The difference is that matrix factorization uses simple inner product for action estimation while neural network employs rather complex multi-layer designs. Despite neural network’s deep structure, traditional designs can not directly model the inner product as in matrix factorization. To combine their advantages, we employ the Product-based Neural Network (PNN, [25]). Compared to traditional neural networks, PNN further introduces a product layer to capture strong interactions such as inner product between embeddings.

The detailed design for item-based site is depicted in Fig. 2(a). We first use one-hot encoding for both user and item as input, then multiply them by user embedding and item embedding matrices  $U^{\mathcal{I}} \in \mathbb{R}^{n \times k}$  and  $A^{\mathcal{I}} \in \mathbb{R}^{m \times k}$  respectively for the embedding. A constant 1 is added here as an additional field. For product layer, we calculate outer product between each pair of embeddings. After that, we append two fully connected hidden layers, with rectified linear unit (relu,  $relu(x) = \max(0, x)$ ) as activation function. For final output layer, we use sigmoid function to model the probability of this action. Formally, we have:

$$\begin{aligned}
 e_p^u, e_q^i &= \text{embeddings for user } p \text{ and item } q \\
 h_0 &= \text{concat}(e_p^u, e_q^i, e_p^u \odot e_q^i) \\
 h_k &= \text{relu}(W_k h_{k-1} + b_k) \\
 \text{output} &= \sigma(h_n W_o + b_o)
 \end{aligned}
 \tag{1}$$

where  $x \odot y$  indicates the flattened outer product of  $x$  and  $y$ ,  $W_k$  and  $b_k$  are the parameters for the fully connected hidden layers.

**Text-Based Site ( $\mathcal{T}$ ):** In microblogging sites, the user actions can be represented by  $(i, \{w_k\})$  tuples indicating that user  $i$  posted or retweeted a microblog



with content  $\{w_k\}$ , where  $w_k$  is the  $k^{th}$  word. Traditionally, topic models such as Latent Dirichlet Allocation [2] are widely used for textual inputs. As neural network develops, word embedding [22] has become the new tool.

We illustrate our model for text-based site in Fig. 2(b). For users, we use one-hot input followed by an embedding layer. For microblogs, we use recurrent neural network with long short-term memory cells to capture the microblog embedding after word embedding layer [24]. Then, we append similar PNN layers as in item-based site for product layer and hidden layers. Finally, sigmoid activation is used for probability estimation.

### 4.2 Modeling the Cross-Site Preferences Consistencies

Most traditional approaches transfer knowledge between sites based on heuristic assumptions and put hard constraints on preferences in different domains. As previously stated, hard transfers do not match with reality.

Instead, we propose JUN based on the assumption that user preferences contain information to reveal the alignment (judging whether two accounts are held by same natural person). Normally, this assumption indicates that we can reveal the alignment by mining user preferences. Because in this task the alignment is given and the goal is to improve user modelings, we reverse the learning direction and further fine-tune the user embeddings using the given alignment.

Specifically, we design a neural network to classify whether two accounts in different sites are held by same natural person, with corresponding user embeddings and side informations as input evidences (showed in Fig. 3). Similar as previous, we use PNN network structure. For side informations, we include username similarity and social ties. Instead of using user embeddings as constant input features, we let them be the same embedding layers as in site-specific action models (Fig. 2), using shared parameters. Then, with the given alignment as supervise training, we can ‘push’ these information back into user embeddings using back propagation, forming a fine-tuning process. Similar methodology has also been widely used for the fine-tuning of word embeddings.

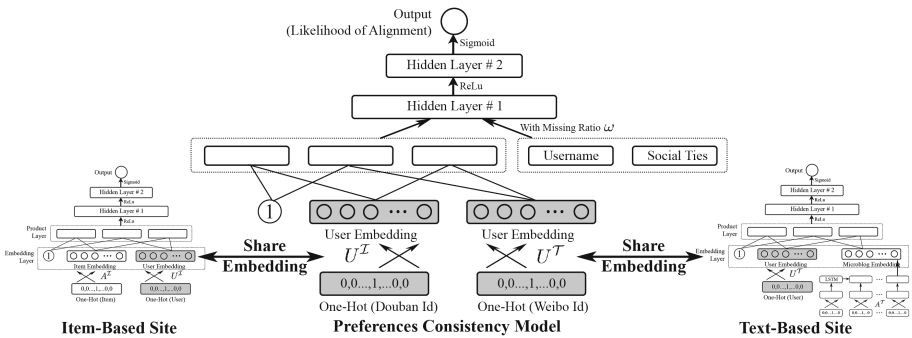


Fig. 3. Fine tuning the user embeddings using alignment information

### 4.3 Learning Techniques

As there are multiple neural networks in JUN, how to conduct the training to achieve expected outcome is also an interesting question.

We train the neural networks by the followings. We first conduct pre-training over site-specific models to initialize user embeddings to represent user preferences. Then, we update the network parameters for preferences consistency model by using user embeddings as constant features. Finally, we update all these networks simultaneously (round-robin, one batch for each network) to conduct the fine-tuning. With this training process, we force the user embeddings to represent user preferences instead of capturing only alignment information.

In order to successfully conduct the fine-tuning using the alignment information, we need to make sure the classifier for alignment does not fully rely on side informations. This could happen because side informations provide strong signals, while user preferences is rather noisy. To prevent this, we simulate information missing on side informations. Specifically, we randomly hide side informations from the model during training with probability  $\omega$ .

### 4.4 Discussion

Comparing to traditional cross-domain or joint user modeling techniques, the major advantage of JUN is its generality in the following aspects:

- **Heterogeneous Actions.** Although in this work we only apply JUN for item-based and text-based sites, JUN’s capability is not limited to these two. It can be easily adopted for any type of action by plugging in corresponding user-embedding based site-specific model as in Sect. 4.1.
- **Employ Other Site-Specific Models.** Similarly, JUN can easily adopt other user embedding-based models to capture the site-specific preferences besides the models proposed in Sect. 4.1.
- **Requirement on Alignment.** Existing cross-site works only work with perfect alignment (bijection) between sites. However, in reality this never happens. With JUN, we only have minimum requirements on alignment. Besides full alignment, JUN can also leverage partial alignment, multi-to-multi alignment, and pairwise score-based alignment as auxiliary data.
- **Extend to Multiple Sites.** JUN can be easily extended for more than two sites by having multiple preferences consistency models, one for each alignment among them. We only evaluate for two sites due to dataset limitation.

Besides generality, JUN is also of high utility value. Most existing recommender systems employ user embedding, or latent vector representation, which both can be easily integrated with JUN. Therefore, JUN can benefit real world applications instead of only for research purpose.

The additional computational cost of JUN is also acceptable. For existing recommender systems, the cost is proportional to the number of user actions. When integrated with JUN, these costs remain (for site-specific preferences models).

The extra computational cost required is for the preferences consistency model. The cost is proportional to the size of alignment, which is same scale for number of users. Comparing to number of user actions, such cost is actually ignorable. Therefore, JUN does not raise a computational cost concern when integrating.

## 5 Experiments

### 5.1 Experimental Settings

**Dataset.** We carry out the experiments using data from Weibo and Douban. The details are previously discussed in Sect. 3.1.

**Methodology.** As there is no ground truth for user preferences, we can only evaluate user modeling through recommender systems. Specifically, we implement recommender system using JUN or existing techniques as user modeling, and then compare the performances of these recommender systems. For Douban and Weibo, the recommender systems rank potential movies and microblogs respectively, based on the estimated likelihoods of action. Area Under the Curve (AUC) is used as the metric to evaluate the ranking.

**Comparing Algorithms.** For comprehensive evaluation, we compare existing approaches for both within-site and cross-site(domain) models. For traditional within-site models, we compare the followings (where  $\mathcal{I}$  and  $\mathcal{T}$  indicate item-based and text-based site respectively):

- ( $\mathcal{I}$ ) **PMF**: Probabilistic Matrix Factorization [23].
- ( $\mathcal{I}$ ) **SVD++**: Matrix factorization with implicit feedback [14].
- ( $\mathcal{I}$ ) **PNN-I**: The isolated site-specific model for item site (Fig. 2(a)).
- ( $\mathcal{T}$ ) **LDA**: Latent Dirichlet Allocation, a widely used topic model [2].
- ( $\mathcal{T}$ ) **PNN-T**: The isolated site-specific model for text site (Fig. 2(b)).

For cross-domain and joint user modeling techniques, we compare:

- ( $\mathcal{I}, \mathcal{T}$ ) **mmTM**: Multi-modal topic model (Fig. 1(c)).
- ( $\mathcal{I}$ ) **TMF**: Topic-Based Matrix Factorization, explicitly adding user’s topic distribution in Weibo as features into feature-based matrix factorization.
- ( $\mathcal{I}, \mathcal{T}$ ) **JUMA**: A graphical model-based joint user modeling approach [4].
- ( $\mathcal{I}, \mathcal{T}$ ) **JUN**: Our approach proposed in this paper.

Although there are plenty cross-domain recommendation techniques, most of them can not be applied under this experiment setting. As we discussed in Sect. 2, most works focus on homogeneous actions, thus can not transfer between microblogging site and movie rating site. There are works modeling between text-based and item-based actions [21], however they require aligned actions for the model learning. For ComSoc in [33], it focuses on borrowing social relations instead of user actions thus is also not suitable for comparison.

**Implementation and Parameters.** We implement JUN using TensorFlow<sup>3</sup>, and use AdamOptimizer for the learning. Unless indicated otherwise, the embedding dimension is 32 for all users, items and words. For hidden layers #1 and #2, we have 100 and 50 hidden units respectively.  $\omega$  is set to 0.5. 80% user actions are used for training. For more details, please refer to our project website<sup>4</sup>.

## 5.2 Performance Comparison

We evaluate aforementioned methods given full alignment as auxiliary information, with training ratio from 40% to 80%. We show the results in Table 1.

For item-based site Douban, JUN achieves better performance comparing to both within-site approaches (PMF, SVD++ and PNN-I) and cross-site approaches (TMF, mmTM, JUMA). Specifically, our approach JUN achieves an AUC of 0.9108 and 0.8984 when training ratio is 80% and 40% respectively, while the best within-site recommender system (PNN-T) achieves only 0.8869 and 0.8695. The relative improvement is 2.69% and 3.32% respectively. Comparing to best cross-site approach, JUN also achieves an additional 2.15% improvement.

For text-based site Weibo, JUN also outperforms all comparing approaches. It achieves an AUC of 0.7443 and 0.7312 under training ratio of 80% and 40%, and the corresponding relative improvement comparing to best existing approach is 2.37% and 3.94% respectively.

Note that for both sites, the performance drop of JUN when reducing training ratio from 80% to 40% is smaller comparing to existing models, indicating that JUN is rather robust to training ratio thanks to the successful data enrichment.

These results indicate that JUN out-performs existing user modeling techniques including within-site and cross-site ones. Also, the improvement occurs on both item-based site and text-based site, indicating that JUN can successfully transfer knowledge between site in both directions. By these experiments, we show that both PNN-I (Fig. 2(a)) and PNN-T (Fig. 2(b)) can be successfully integrated with JUN for a further improvement. For other unevaluated embedding-based techniques, we believe that similar improvement can also be achieved.

## 5.3 Cold-Start Scenario

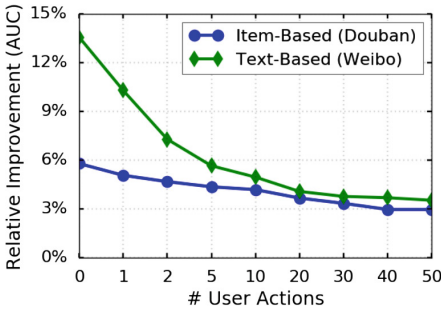
We simulate cold-start scenarios by limiting the number of user actions used for training. We depict the relative AUC improvements over users with different number of training actions in Fig. 4(a). Results in both sites indicate that performance improvement is much higher when dealing with cold users comparing to non-cold users. We achieve a relative improvement of 13.54% for users with no historical actions in Weibo and 5.77% in Douban, indicating that JUN succeeded in leveraging cross-site actions for cold-start problem. Note that the improvement is more significant in text-based site comparing to item-based site. This

<sup>3</sup> <https://www.tensorflow.org/>.

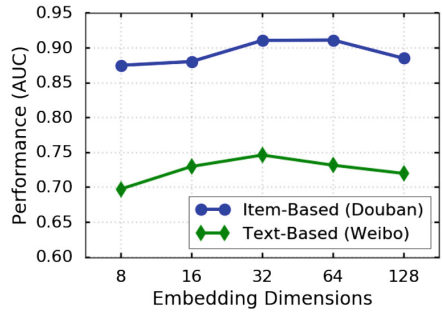
<sup>4</sup> Implementation and dataset are available at: <http://www.apexlab.org/projects/32>.

**Table 1.** Experimental results, varying training ratio

Target	Type	ALGS	AUC score, varying training ratio		
			0.4	0.6	0.8
Item-based (Douban)	Within site	PMF	0.7969 $\pm$ 0.0017	0.8087 $\pm$ 0.0016	0.8196 $\pm$ 0.0015
		SVD++	0.8598 $\pm$ 0.0013	0.8738 $\pm$ 0.0016	0.8849 $\pm$ 0.0014
		PNN-I	0.8695 $\pm$ 0.0016	0.8787 $\pm$ 0.0013	0.8869 $\pm$ 0.0012
	Cross site	mmTM	0.7494 $\pm$ 0.0019	0.7510 $\pm$ 0.0014	0.7608 $\pm$ 0.0012
		TMF	0.8514 $\pm$ 0.0016	0.8654 $\pm$ 0.0013	0.8763 $\pm$ 0.0015
		JUMA	0.8790 $\pm$ 0.0018	0.8889 $\pm$ 0.0014	0.8916 $\pm$ 0.0015
	<b>JUN</b>	<b>0.8984 <math>\pm</math> 0.0015</b>	<b>0.9052 <math>\pm</math> 0.0013</b>	<b>0.9108 <math>\pm</math> 0.0013</b>	
Text-Based (Weibo)	Within site	LDA	0.6514 $\pm$ 0.0017	0.6694 $\pm$ 0.0016	0.6839 $\pm$ 0.0015
		PNN-T	0.7035 $\pm$ 0.0018	0.7129 $\pm$ 0.0015	0.7271 $\pm$ 0.0013
	Cross site	mmTM	0.6652 $\pm$ 0.0018	0.6776 $\pm$ 0.0015	0.6828 $\pm$ 0.0016
		JUMA	0.6824 $\pm$ 0.0014	0.6976 $\pm$ 0.0014	0.7120 $\pm$ 0.0012
		<b>JUN</b>	<b>0.7312 <math>\pm</math> 0.0016</b>	<b>0.7384 <math>\pm</math> 0.0013</b>	<b>0.7443 <math>\pm</math> 0.0014</b>



(a) Cold Start Scenarios



(b) Performance vs Dimension Size

**Fig. 4.** Detailed analysis of JUN.

is because non-personalized recommendation still works well in item-based scenario by directly ranking the items by popularity, but not in text-based scenario (Fig. 4).

#### 5.4 Parameter Tunning

We vary the embedding dimensions of JUN for detailed analysis. We evaluate using dimension size from 8 to 128 and report the results in Fig. 4(b). According to the results, embedding of 32 dimensions gives the best overall performance. Large dimension size leads to over fitting and higher computational cost. Also note that when reducing embedding dimensions, the performance drop for item-based site is smaller than for text-based site. This may also due to the quality of non-personalized recommender system in item-based site is rather good.

Experiments also indicate that performance is not very sensitive to other parameters besides embedding dimension.

## 5.5 Partial Alignment and Score-Based Alignment

A major limitation of existing joint user modeling or cross-domain user modeling technique is that they can only be applied to fully aligned sites. However, such situation only exist in ideal research environment but not the reality. To demonstrate that JUN can also be applied to partial alignment and score-based alignment generated using network aligners, we conduct experiments for these scenarios. For score-based alignment, we use BASS as the network aligner [3].

We show the results in Table 2. For this set of experiments, we report not only the overall AUC for all users, but also the average AUC for aligned users and unaligned users respectively to show the effect of JUN for different user groups according to the alignment. The results match with our expectation that full alignment results in the most significant improvement according to overall AUC, and the improvement of partial alignment lies between full alignment and no alignment depending on the alignment ratio. For aligned users, the improvements are mostly consistent for all alignment methods. Note that for unaligned users, JUN also achieves a slightly improvement. This may due to the knowledge transfer between aligned users leads to better embedding spaces for both user and items (or words, etc.), thus also improve the quality for unaligned users.

**Table 2.** Apply JUN for partial alignment and score-based alignment

Alignment	Aligned users		Unaligned users		All users	
	$\mathcal{I}$	$\mathcal{T}$	$\mathcal{I}$	$\mathcal{T}$	$\mathcal{I}$	$\mathcal{T}$
No alignment	n/a	n/a	0.8869	0.7271	0.8869	0.7271
40% alignment	0.9086	0.7425	0.8931	0.7342	0.8993	0.7375
80% alignment	0.9113	0.7438	0.8964	0.7393	0.9083	0.7429
Full alignment	0.9108	0.7443	n/a	n/a	0.9108	0.7443
Score-based alignment	n/a	n/a	n/a	n/a	0.9086	0.7419

## 6 Conclusion and Future Works

In this paper, we aim at improving the quality of user modeling by conducting joint user modeling across aligned heterogeneous sites using neural networks. To overcome the limitations of existing cross-domain and joint user modeling works regarding the modeling of heterogeneous actions, the requirement of full alignment and the design of coupling strength, we propose a novel neural network-based framework JUN to tackle this task. JUN takes advantage of neural network’s capability for capturing heterogeneous data and its ability for mining high-level non-linear correlations. Comparing to existing works, JUN can be further applied for scenarios where only partial alignment or pairwise score-based alignment are available. Also, JUN models preferences consistency using fine-tuning neural network instead of hard constraints, leading to better knowledge transfer across sites. We conduct extensive experiments using real data from

Douban and Weibo to evaluate JUN's performance. Results indicate that JUN outperforms existing works in both sites, and successfully alleviates the cold-start problem. We achieve relative improvement of 2.96% and 2.37% for item-based and text-based sites respectively. For cold-start scenarios, we achieve relative improvement of 5.77% and 13.54% respectively.

For future works, we may consider integrating social relations from the aligned sites into JUN. Also, we are interested in integrating JUN with network aligners to conduct the aligning and user modeling simultaneous as these tasks are high related with each other.

## References

1. Berkovsky, S., Kuflik, T., Ricci, F.: Cross-domain mediation in collaborative filtering. In: Conati, C., McCoy, K., Paliouras, G. (eds.) UM 2007. LNCS (LNAI), vol. 4511, pp. 355–359. Springer, Heidelberg (2007). [https://doi.org/10.1007/978-3-540-73078-1\\_44](https://doi.org/10.1007/978-3-540-73078-1_44)
2. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. *JML* **3**, 993–1022 (2003)
3. Cao, X., Yu, Y.: BASS: a bootstrapping approach for aligning heterogenous social networks. In: Frasconi, P., Landwehr, N., Manco, G., Vreeken, J. (eds.) ECML PKDD 2016. LNCS (LNAI), vol. 9851, pp. 459–475. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46128-1\\_29](https://doi.org/10.1007/978-3-319-46128-1_29)
4. Cao, X., Yu, Y.: Joint user modeling across aligned heterogeneous sites. In: Proceedings of the 10th ACM Conference on Recommender Systems, pp. 83–90. ACM (2016)
5. Cheng, H.T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhya, H., Anderson, G., Corrado, G., Chai, W., Ispir, M., et al.: Wide & deep learning for recommender systems. In: Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, pp. 7–10. ACM (2016)
6. Fernández-Tobías, I., Cantador, I., Kaminskas, M., Ricci, F.: A generic semantic-based framework for cross-domain recommendation. In: Proceedings of the 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems, pp. 25–32. ACM (2011)
7. Fernández-Tobías, I., Cantador, I., Kaminskas, M., Ricci, F.: Cross-domain recommender systems: a survey of the state of the art. In: SCIR (2012)
8. Gao, H., Tang, J., Liu, H.: Addressing the cold-start problem in location recommendation using geo-social correlations. *Data Min. Knowl. Discov.* **29**, 1–25 (2014)
9. Godin, F., Slavkovikj, V., De Neve, W., Schrauwen, B., Van de Walle, R.: Using topic models for Twitter hashtag recommendation. In: WWW (2013)
10. Golbandi, N., Koren, Y., Lempel, R.: Adaptive bootstrapping of recommender systems using decision trees. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, pp. 595–604. ACM (2011)
11. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., Chua, T.S.: Neural collaborative filtering. In: Proceedings of the 26th International World Wide Web Conference (2017)
12. Iofciu, T., Fankhauser, P., Abel, F., Bischoff, K.: Identifying users across social tagging systems. In: ICWSM (2011)

13. Koenigstein, N., Dror, G., Koren, Y.: Yahoo! music recommendations: modeling music ratings with temporal dynamics and item taxonomy. In: Proceedings of the Fifth ACM Conference on Recommender Systems, pp. 165–172. ACM (2011)
14. Koren, Y.: Factorization meets the neighborhood: a multifaceted collaborative filtering model. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 426–434. ACM (2008)
15. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. *Computer* **42**, 30–37 (2009)
16. Lam, X.N., Vu, T., Le, T.D., Duong, A.D.: Addressing cold-start problem in recommendation systems. In: Proceedings of the 2nd International Conference on Ubiquitous Information Management and Communication, pp. 208–211. ACM (2008)
17. Levandoski, J.J., Sarwat, M., Eldawy, A., Mokbel, M.F.: Lars: a location-aware recommender system. In: 2012 IEEE 28th International Conference on Data Engineering (ICDE), pp. 450–461. IEEE (2012)
18. Liu, H., Maes, P.: Interestmap: harvesting social network profiles for recommendations. *Beyond Personalization-IUI*, p. 56 (2005)
19. Liu, J., Zhang, F., Song, X., Song, Y.L., Lin, C.Y., Hon, H.W.: What’s in a name? An unsupervised approach to link users across communities. In: Proceedings of the 6th ACM International Conference on Web Search and Data Mining, pp. 495–504. ACM (2013)
20. Liu, S., Wang, S., Zhu, F., Zhang, J., Krishnan, R.: Hydra: large-scale social identity linkage via heterogeneous behavior modeling. In: SIGMOD (2014)
21. McAuley, J., Leskovec, J.: Hidden factors and hidden topics: understanding rating dimensions with review text. In: RecSys, pp. 165–172. ACM (2013)
22. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint [arXiv:1301.3781](https://arxiv.org/abs/1301.3781) (2013)
23. Mnih, A., Salakhutdinov, R.: Probabilistic matrix factorization. In: NIPS, pp. 1257–1264 (2007)
24. Palangi, H., Deng, L., Shen, Y., Gao, J., He, X., Chen, J., Song, X., Ward, R.: Deep sentence embedding using long short-term memory networks: analysis and application to information retrieval. *IEEE/ACM Trans. Audio Speech Lang. Process. (TASLP)* **24**(4), 694–707 (2016)
25. Qu, Y., Cai, H., Ren, K., Zhang, W., Yu, Y., Wen, Y., Wang, J.: Product-based neural networks for user response prediction. arXiv preprint [arXiv:1611.00144](https://arxiv.org/abs/1611.00144) (2016)
26. Sahebi, S., Brusilovsky, P.: Cross-domain collaborative recommendation in a cold-start context: the impact of user profile size on the quality of recommendation. In: Carberry, S., Weibelzahl, S., Micarelli, A., Semeraro, G. (eds.) UMAP 2013. LNCS, vol. 7899, pp. 289–295. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-38844-6\\_25](https://doi.org/10.1007/978-3-642-38844-6_25)
27. Schein, A.I., Popescul, A., Ungar, L.H., Pennock, D.M.: Methods and metrics for cold-start recommendations. In: SIGIR, pp. 253–260 (2002)
28. Tan, S., Bu, J., Qin, X., Chen, C., Cai, D.: Cross domain recommendation based on multi-type media fusion. *Neurocomputing* **127**, 124–134 (2014)
29. Tan, S., Guan, Z., Cai, D., Qin, X., Bu, J., Chen, C.: Mapping users across networks by manifold alignment on hypergraph. In: AAAI (2014)
30. Vasile, F., Smirnova, E., Conneau, A.: Meta-Prod2Vec: product embeddings using side-information for recommendation. In: Proceedings of the 10th ACM Conference on Recommender Systems, pp. 225–232. ACM (2016)
31. Zhang, J., Kong, X., Yu, P.S.: Predicting social links for new users across aligned heterogeneous social networks. In: 2013 IEEE 13th International Conference on Data Mining, pp. 1289–1294. IEEE (2013)



32. Zhang, Z.K., Liu, C., Zhang, Y.C., Zhou, T.: Solving the cold-start problem in recommender systems with social tags. *EPL* **92**(2), Article no. 28002 (2010)
33. Zhong, E., Fan, W., Wang, J., Xiao, L., Li, Y.: ComSoc: adaptive transfer of user behaviors over composite social network. In: *Proceedings of the 18th ACM SIGKDD*, pp. 696–704. ACM (2012)
34. Zhou, K., Yang, S.H., Zha, H.: Functional matrix factorizations for cold-start recommendation. In: *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 315–324. ACM (2011)