Dynamic Traffic Network Based Multi-Modal Travel Mode Fusion Recommendation

Nannan Jia[1], Mengmeng Chang[1], Zhiming Ding[2](✉), Zunhao Liu[1], Bowen Yang[1], Lei Yuan[1], and Lutong Li[1]

[1] Beijing University of Technology, Beijing 100124, China
{jianannan, changmengmeng, bovin.y, yuanlei, lilutong}@emails.bjut.edu.cn
[2] Institute of Software, Chinese Academy of Sciences, Beijing 100190, China
zhiming@iscas.ac.cn

Abstract. The travel problem is a challenge to the city's overall development. With the increase of the number of cars and the increase of urban population density, intelligent travel mode provides new solutions to solve these problems. However, the existing research on the choice of travel modes for residents only considers the current traffic conditions, and the preferences of individual users for travel modes are poorly considered, which cannot meet the personalized travel needs of users. From this perspective, a heterogeneous information network based on users' spatial-temporal travel trajectories is proposed in this paper. Considering the dynamic traffic network that is constantly changing during the travel process, and using the graph neural network guided by the meta-path to dynamically model the user and travel mode. Features embedding with rich interactive information, so as to fully learn the users' preferences for travel modes in the time-space travel trajectory, and recommend travel modes that meet personalized needs to users. Finally, the effectiveness of the proposed method is demonstrated by experimental evaluation on real-world datasets.

Keywords: Heterogeneous information network · Graph neural network · Meta-path · Multi-modal travel mode · Fusion recommendation

1 Introduction

With the continuous development of urbanization and the improvement of economic level, people's travel demands become increasingly diverse. When demand exceeds the carrying capacity of the transport system, various traffic and even environmental problems will arise. The complex selection behavior of people in multi-mode transportation network [26] determines the distribution of travel demand on the transportation network. From the perspective of balancing the distribution of traffic demand [3], personalized travel mode recommendation has great research value.

Nowadays, in the era of big data, as an information filtering technology, recommendation system is particularly important for users to recommend information that meets their individual needs. Heterogeneous information networks [4, 5] consider different types of objects, different connection relationships between objects, and

H. Gao and X. Wang (Eds.): CollaborateCom 2021, LNICST 406, pp. 299–317, 2021.
https://doi.org/10.1007/978-3-030-92635-9_18

attribute information, fully reflecting the interaction between different objects in the recommendation system, from which implicit information is mined and deeper regularities are learned, greatly improving recommendation accuracy and opening up a new path for personalized recommendation technology. Besides, it can be seen that traditional user-item recommendations are themselves heterogeneous bipartite graphs, so the use of HIN (Heterogeneous information networks, HIN) in recommendation techniques is inevitable. Common heterogeneous information networks such as social networks, biological neural networks, academic networks, etc., but there is more to it than that - the real world is rich in data, and information networks are everywhere. For example, construct a heterogeneous information network based on the types of movies, directors, and user attributes to express user preferences in a more granular manner [15, 16]. There are many types of research on recommendation by constructing heterogeneous information networks, such as malicious account detection [17] and search intent recommendation [10]. This paper finds that the travel behavior of users on a multi-mode transportation network is also an embodiment of information network, and then proposed to apply the representation learning technology in heterogeneous information network to transportation travel, and combined with dynamic transportation network to recommend multi-mode travel mode.

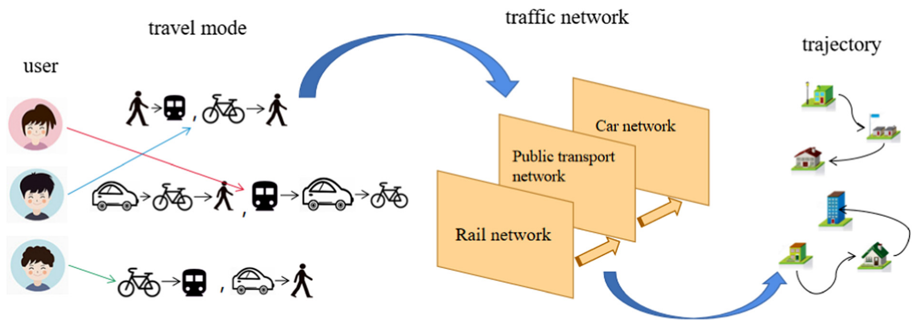


Fig. 1. Interaction of users' Spatial-temporal travel trajectories and transport networks.

Figure 1 illustrates the interaction between the user's Spatial-temporal travel trajectory and the transport network, where the user uses multiple modes of travel to continuously shift between this complex transport network, eventually forming a chain of trajectories. Among them, the complex transportation network is composed of multimodal travel modes. The travel mode is the basic attribute of understanding the user's mobile behavior, and the transportation mode used by the user in the trajectory can reflect the regularity of the user's mobile state. Therefore, most scholars have always been concerned about travel Research fields related to methods have attracted much attention. YuZheng [18–20] et al. used methods such as change-point segmentation based on the identification of travel patterns in GPS trajectories to understand users' travel patterns. Dabiri S [21] et al. used convolutional neural networks to train models for automatic inference of travel patterns. Yao [1] et al. proposed a travel mode choice model that considers the combined effects of rational decision making and inherent choice preferences to analyze users' choice of travel mode [6, 13]. Used multiple sources of data to build a model of residents' travel mode choice and to analyze their travel mode choice behavior. This paper proposes to apply heterogeneous information network representation learning to transport travel to complete personalized recommendations of travel modes to fill the gap to a certain extent. Gunjan Kumar [2] et al. used a sequence-based approach to personalize travel style recommendations to users, but it did not sufficiently take into account the contextual information of the recommended objects and the interaction between them, which is well remedied by our study. Personalized recommendation of travel modes remains an underexplored task, with the following main contributions:

1. Considering the more abundant characteristics of learning users and modes of travel, we innovatively propose building heterogeneous information networks with users, modes of travel, Spatial-temporal attributes (e.g. time, etc.) strongly related to modes of travel, and starting and ending locations.
2. Using the Graphic Neural Network guided by the meta-path to dynamically model users and travel modes, we can get the embedded features with rich interactive information, and fully learn user's preferences for travel modes from HIN.
3. The dynamic traffic network that changes with time during the trip is considered to increase the probability of selecting more time-saving and energy-saving trip modes. Finally, the feasibility and effectiveness of the proposed method are verified by real datasets.

The remainder of this paper is structured as follows. In Sect. 2 some basic concepts used in the article are introduced. Section 3 presents our fusion recommendation model and the main techniques used. Section 4 presents a comprehensive evaluation of our model on a real dataset. Finally, conclusions are drawn in Sect. 5.

2 Concepts Used in the Paper

In this section, we introduce the heterogeneous information networks relevant to the fused recommendation content of this paper and some basic concepts used in the recommendation model.

2.1 Definition of the Fusion Recommendation Problem

Define a set $M = \{U, V, T, L\}$, where U denotes the set of users; V denotes the travel mode used by the user in the travel trajectory, mainly including metro, bus, driving, taxi, since car, walking; T denotes the start time when the user uses a certain travel mode; L denotes the starting and ending location of the user at a certain time when using a certain travel mode; all these belong to the node type in the heterogeneous transportation travel network. By analyzing the user's historical travel trajectory and learning the user's personalized preference for travel mode, combined with the impact of dynamic traffic network changes over time on travel mode, this fusion recommendation finally recommends more time-saving and energy-saving multimodal travel mode $v \in V$ in line with personal preference for users $u \in U$.

2.2 Heterogeneous Transport Travel Networks

In a multi-modal transport network, users take different modes of travel to reach their destination and there is a certain connection between them. In this paper, we construct a heterogeneous information network (see Fig. 2(a)) with the temporal attributes of a user, travel mode, origin and destination, and travel mode to obtain more potential features by learning the semantic relationships between objects such as user and user, user and travel mode, travel mode and its corresponding origin and destination (note: at this point, the temporal attributes are closely connected to the user), and to fully express the user's personalized preference for travel mode. To enable a better understanding of heterogeneous information networks, the construction of a network schema [4] is shown in Fig. 2(b), which specifies the type constraints on the set of objects and the connections between them: there is a connection between users and travel modes, indicating a use-and-used relationship; there is a connection between locations and travel modes, indicating a guide-and-directed relationship, etc.

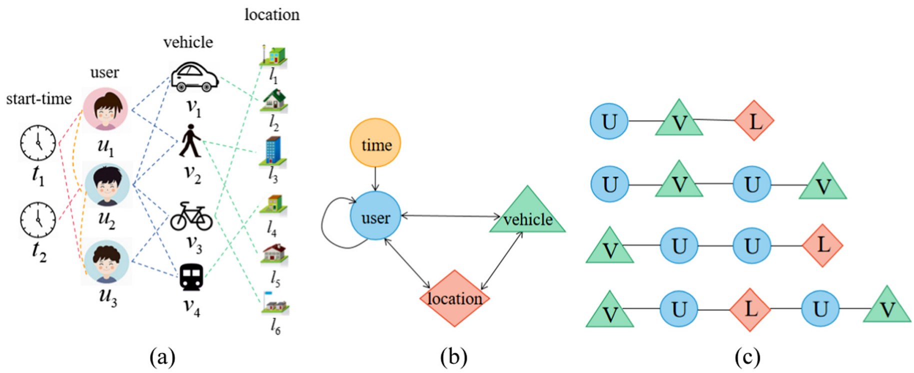


Fig. 2. Interaction of users' Spatial-temporal travel trajectories and transport networks.

2.3 Meta-paths Extraction Based on User Trajectory

A meta-path [7] is a path defined over a network schema linking two types of objects, representing a composite relationship between object types, and is used to extract information about the interaction between objects. Figure 2(c) shows the four meta-paths used in this paper, where user and time are closely linked. The meta-path "user-vehicle-location" indicates that the user has traveled from a certain location to a certain location at a certain time. The meta-path from "user-vehicle-user-vehicle" indicates that a user has used a travel mode at a certain time and that similar users have also used this mode of travel and have also used other travel modes travel. "Vehicle-user-user-location" indicates that a similar user used a travel mode when arriving at a similar starting and ending location; "vehicle-user-location-user-vehicle" indicates that a similar user used a travel mode when arriving at a similar location at a similar time. Meta paths are used in this recommender system to obtain semantic information between objects in a heterogeneous information network.

2.4 Meta-path Guided Neighbors

This paper focuses on integrating rich information through aggregated meta-path-guided neighborhoods. A detailed description of the meta-path-guided neighborhood, as exemplified by the user u_2 in Fig. 2(a): The first-order neighbors of user u_2 under the meta-path $\Phi : U - V - L$ are that $N^1_{u_2} = \{v_1, v_2, v_3, v_4\}$, The second-order neighbors of u_2 are the first-order neighbors of all nodes in the first-order neighborhood, so its second-order neighbors are $N^2_{u_2} = \{l_1, l_2, l_3, l_4, l_5, l_6\}$. Therefore all neighbors of the user u_2 are denoted as $N_{u_2} = \left\{N^0_{u_2}, N^1_{u_2}, N^2_{u_2}\right\} = \{u_2, v_1, v_2, v_3, v_4, l_1, l_2, l_3, l_4, l_5, l_6\}$.

3 Heterogeneous Transport Travel Network Recommendation Model

In this section, we will introduce in detail the multi-modal travel mode fusion recommendation model (DTN-MMTMRec) based on meta-path under the condition of dynamic transportation network. The idea of this fusion recommendation is to construct the user's Spatial-temporal travel trajectory into a heterogeneous information network, and select different meta-paths, using the different semantic information they express, to build a graph neural network, and to enrich the node embedding of the user and travel mode by aggregating the information of neighboring nodes guided by the meta-paths, i.e. to better learn the representation of the user and travel mode, and to recommend a travel mode for the user that meets personalized needs. Figure 3 illustrates the overall framework of this fusion recommendation model. Heterogeneous information network as input, the embedding layer preprocesses the model data with content features to generate an initial node embedding. Then, in the meta-path aggregation layer, the selected meta-path is aggregated with its guided neighbor nodes for dynamic modeling, i.e. the semantic information between the target node and its neighbors is aggregated to obtain rich embeddings of users and travel modes, and finally, the feature embeddings of users and travel modes are fused for multimodal travel mode recommendation. The application of our model to travel mode recommendations is expected to yield better performance, and the use of meta-paths will also improve the interpretability of the recommendation results. The details are described in detail in the following subsections.

3.1 Initial Embedding

Following previous work [7, 8], different types of node features in heterogeneous graphs were mapped to the same vector space, applying a specific type of linear transformation to each type of node:

$$E_v = W_A \cdot x^A_v, E^0_v = x_v \tag{1}$$

where $x_v \in R^d$ is the original feature vector, $E_v \in R^{d'}$ is the projection latent vector of node v. $W_A \in R^{d' \times d}$ is the parameter weight matrix of type A nodes. After applying this

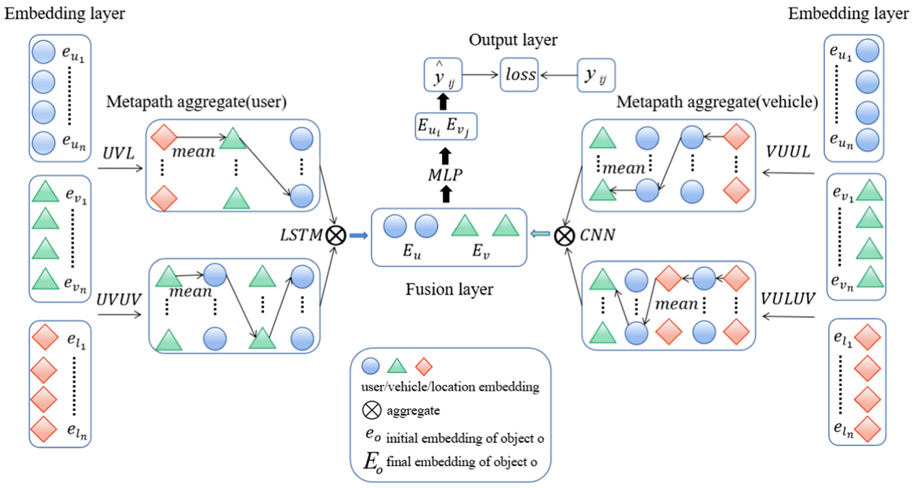


Fig. 3. Overall framework for fusion recommendations.

operation, the projected features of all nodes have the same dimensionality, facilitating the aggregation process for the next model component.

3.2 Practice of Meta-path

An important step in existing recommender systems is to model the interactions between objects to learn more information about the characteristics of the target object. For recommender systems with only two objects, user and item, traditional methods such as matrix decomposition and the more popular [8] method of obtaining higher-order semantics to enrich object embeddings based on a two-part interaction graph of user-item have achieved good results in recent years. For heterogeneous graphs consisting of multiple types of nodes and multiple types of edges, traditional learning embedding methods only involve nodes and do not capture information about neighbouring nodes and the interaction information between neighbouring nodes. With the widespread use of graph neural networks, this deficiency is well filled, where a fixed size number of neighbouring nodes is selected using random wandering for heterogeneous graphs in [11], and then a suitable aggregation function is selected to aggregate the neighbouring vertex feature information to the target node. Similarly, this paper uses different multi-hop meta-paths to determine the neighbors of the target node and aggregates the feature information of the neighboring nodes to derive the final embedding of the target node.

Take Fig. 2(a) as an example to introduce in detail the application of meta-paths in heterogeneous graph neural networks. The feature aggregation process for the target node u_2 is described here. Define two meta-paths $\Phi_{uvl} : U - V - L$, $\Phi_{UVUV} : U - V - U - V$, The initial feature vector of each node is known, and Fig. 4 illustrates the two processes of meta-path aggregation.

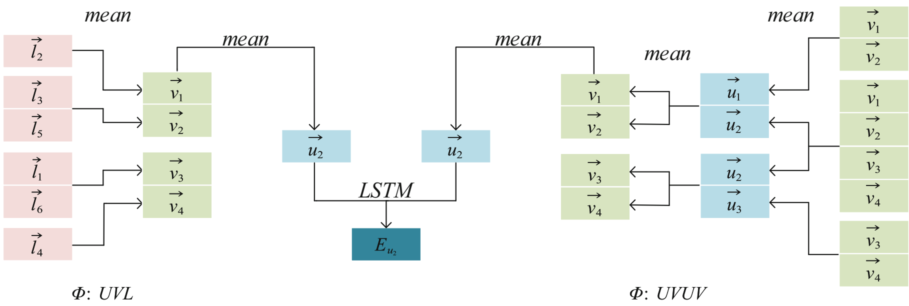


Fig. 4. The meta-path aggregation process.

Internal aggregation of meta paths: choose a suitable aggregation function to aggregate the embeddings of second-order neighbor nodes to obtain the embeddings of first-order neighbor nodes, and then aggregate the embeddings of first-order neighbor nodes to obtain the embeddings of target nodes, as follows:

$$E_{u_2}^{\Phi_{uvl}} = f(e_{v_1}, e_{v_2}, e_{v_3}, e_{v_4}) = f(e_{l_1}, e_{l_2}, e_{l_3}, e_{l_4}, e_{l_5}, e_{l_6}) \tag{2}$$

In formula 2, f is the aggregation function that aggregates the first-order neighbor, second-order neighbor embeddings of u_2 on the specified meta-path. $(e_{l_1}, e_{l_2}, e_{l_3}, e_{l_4}, e_{l_5}, e_{l_6})$ is the second-order neighbor node embedding of user u_2 on the meta-path, $(e_{v_1}, e_{v_2}, e_{v_3}, e_{v_4})$ is the first-order neighbor node embedding, which is updated by aggregating the second-order neighbor node embeddings, and then aggregating the first-order neighbor embeddings to obtain the final embedding of user u_2 on the meta-path Φ_{uvl}.

Aggregation between meta paths: aggregate user u_2 embeddings from different meta paths to get the final embedding of u_2:

$$E_{u_2} = f(E_{u_2}^{\Phi_{uvl}}, E_{u_2}^{\Phi_{uvuv}}, \ldots) \tag{3}$$

3.3 Meta-path Aggregation Functions

In the meta-path aggregation layer, a basic graph-based neural network is used. The basic idea is to embedding nodes based on their local neighbor information [9]. In this paper, a neural network is used to aggregate the information of target nodes and their neighbors [14, 27].

The aggregation process is divided into two steps, aggregation within a single meta-path and aggregation between different meta-paths. The former mainly aggregates the neighboring nodes of a node, and the average aggregation function is used in this paper. The latter is a feature embedding that aggregates different meta-paths. As different meta-paths express different semantics, the average aggregation function is no longer used, and the LSTM aggregation function or CNN aggregation function is chosen here according to the specific data characteristics. The different effects on recommendation performance will also be demonstrated in the subsequent experimental evaluation of the three aggregation functions shown below, where E_u^{k-1} denotes the embedding of node u in layer $k-1$:

Mean Aggregation Function. The eigenvectors of the target node's neighbors are averaged and then stitched together with the target node eigenvectors, with two non-linear transformations in between.

$$AGG = \sum\nolimits_{u \in N(v)} \frac{E_u^{k-1}}{|N(v)|} \tag{4}$$

LSTM Aggregation Function. The neighboring nodes of the target node are randomly disordered as the input sequence, and the resulting vector representation and the vector representation of the target node are separately stitched together after a non-linear transformation to obtain the vector representation of the target node at that layer.

$$AGG = LSTM\left(\left[E_u^{k-1}, \forall u \in N(v)\right]\right) \tag{5}$$

CNN Aggregation Function. Unlike the LSTM aggregation function, it does not have to consider time series conditions and can be used directly as an input sequence to generate new features by convolution operations.

$$AGG = CNN\left(\left[E_u^{k-1}, \forall u \in N(v)\right]\right) \tag{6}$$

3.4 Semantic Aggregation

The fusion recommendation model in this paper updates the features of nodes in heterogeneous information networks by aggregating neighbors guided by different meta paths to obtain node embedding with rich semantics. In this section, the process of embedding users and modes of travel as target nodes are described in detail.

As represented by the meta-path aggregation layer in Fig. 3. Firstly, a user is regarded as the target node (it can be a multi-hop meta-path or single-hop meta-path), and then the neighbor node features are aggregated one by one according to the selected meta-path. Using the meta-path UVL as an example, the initial embedding of the travel mode and start/stop location is known and the aggregation process is as follows.

- Aggregation between nodes within a single meta-path

The average aggregation function is used to aggregate the embedding of the second-order neighbors, and the aggregation result is spliced with the first-order neighbors to update the embedding of the first-order neighbors:

$$E_{v_j}^{\Phi_{uvl}} = f(E_{l_1}^{\Phi_{uvl}}, E_{l_2}^{\Phi_{uvl}}, E_{l_3}^{\Phi_{uvl}}, \ldots) \tag{7}$$

The feature embedding of the target node is updated using the same average aggregation function to aggregate the first-order neighbor node embedding:

$$E_{u_i}^{\Phi_{uvl}} = f(E_{v_1}^{\Phi_{uvl}}, E_{v_2}^{\Phi_{uvl}}, \ldots) \tag{8}$$

- Aggregation between different meta-paths

In this paper's recommendation model, time plays a key role in recommending to users a travel mode that meets their personalized needs, so when modeling users dynamically, users are temporal, and the LSTM aggregation function is used here to better aggregate the embedding of users under different meta-paths. where Φ_i is a different meta-path with the user as the target node:

$$E_{u_i} = f(E_{u_i}^{\Phi_1}, E_{u_i}^{\Phi_2}, E_{u_i}^{\Phi_3}, \ldots, E_{u_i}^{\Phi_k}) \tag{9}$$

The final embedding of all user nodes is obtained by dynamically modeling all users in the above way:

$$\{E_{u_1}, E_{u_2}, E_{u_3}, \ldots, E_{u_n}\} \tag{10}$$

The dynamic modeling of travel mode is similar to this. Based on the different meta-paths with travel mode as the target node, the average aggregation function is first used to aggregate the features of the neighboring nodes within the meta-paths one by one, and then the CNN aggregation function is used to complete the feature aggregation between different meta-paths to obtain the final embedding of all travel mode nodes. where Φ_i is the different meta-paths with travel mode as the target node:

$$E_{v_i} = f(E_{v_i}^{\Phi_1}, E_{v_i}^{\Phi_2}, E_{v_i}^{\Phi_3}, \ldots, E_{v_i}^{\Phi_k}), \{E_{v_1}, E_{v_2}, E_{v_3}, \ldots, E_{v_n}\} \tag{11}$$

3.5 Evaluation Prediction

In this recommendation model, we predict the probability $\widehat{y}_{ij}$ of a user choosing a travel mode at a given time, which is in the range [0, 1]. The final embeddings of users and travel modes are obtained by aggregating meta-path-guided neighbor embeddings, and then we fuse the user, travel mode node embeddings for connectivity and finally input them into the MLP to predict the score $\widehat{y}_{ij}$:

$$\hat{y}_{ij} = sigmoid\left(f\left(U_i \otimes V_j\right)\right) \tag{12}$$

In formula (12), f is the MLP layer with only one output, sigmoid is the activation layer and $\otimes$ is the embedded link operator.

The loss function in this model uses a point-by-point loss function, and the model is adjusted by the loss function to produce optimal results, $\mathcal{Y}$ and $\mathcal{Y}^-$ are the instance sets of positive samples and negative samples respectively:

$$J = \sum\nolimits_{i,j \in \mathcal{Y} \cup \mathcal{Y}^-} \left(y_{ij} \log \hat{y}_{ij} + \left(1 - y_{ij}\right) \log\left(1 - \hat{y}_{ij}\right) \right) \tag{13}$$

4 Experiments and Analysis

We perform experimental evaluations on real-world datasets to verify the effectiveness of the model DTN-MMTMRec on multimodal travel mode recommendations.

Table 1. The selected meta-paths used in dataset and the meaning expressed.

Meta-paths	The meaning expressed
U-V-L	User preferences for travel modes
U-V-U-V	Preferences for travel modes between similar users
V-U-U-L	Add location to express travel preferences at a more granular level
V-U-L-U-V	Travel preferences of users arriving at similar locations

4.1 Dataset

Description of the Dataset. To evaluate the effectiveness of the DTN-MMTMRec model in this paper, we conducted experiments on a dataset from the Microsoft Geolife project [19, 20, 22, 23]. The dataset consists of two parts, one is the user's track log, and each track is a sequence of time stamp points, each of which contains its associated longitude, dimension, time, and so on. The other part is the travel mode label file corresponding to the user's track log, which contains the travel mode and starts time used by the user in the track.

The dataset contains 10 different modes of travel, namely bicycle, bus, car, metro, taxi, train, walking, plane, boat, and running. Table 1 shows the percentage of days that each mode of travel occurred at least once for all users, which shows that some modes of travel were used frequently and some were used occasionally. For example, walking is one of the most frequently used by users, while airplanes, sailing boats, and running are hardly ever used. Besides, we observe that users in the dataset use between two and five different modes of travel per day. This was filtered by retaining data for only six modes of travel - walking, cycling, bus, subway, driving, private car, and taxi - and then fusing the data based on the start time of using a particular mode of travel in the labeled data.

Table 2. Percentage of days each mode of travel occurs at least once among all users.

Travle mode	Percent	Rank
Walk	38%	1
Bus	24%	2
Subway	12%	3
Bike	12%	4
Taxi	7%	5
Car	6%	6
Airplane	0.8%	7
Train	0.2%	8
Boat	–	9
Run	–	10

Data Fusion. The main task of data fusion is to incorporate time attributes with key impacts into the data and establish the interaction among users, modes of travel, and starting and ending locations. In the last section, we will make statistics and analysis of the user's space-time travel trajectory data, and ultimately select the six travel modes commonly used by urban residents. After a series of cleaning, filtering, and integrating data, the final results are as follows: [User, mode of travel, start time, start and end location]. Previously, the user data of the recommendation system only contained the user and then established a relationship with the recommendation object, even if the time attribute was added, the relationship between the user, the time, and the recommendation object was also established. In this paper, the data structure is improved as shown in Fig. 5, combining the user and time together, from the original user using a certain travel mode to a certain user using a certain travel mode at a certain time, which can not only capture the user's travel behavior data more accurately but also refine the data on the basis of the original data, which provides a great advantage to the subsequent model training work. The paper then establishes the interaction between the three objects. The user uses a certain mode of travel at a certain time: $\{u_i - t_j : v_k\}$,

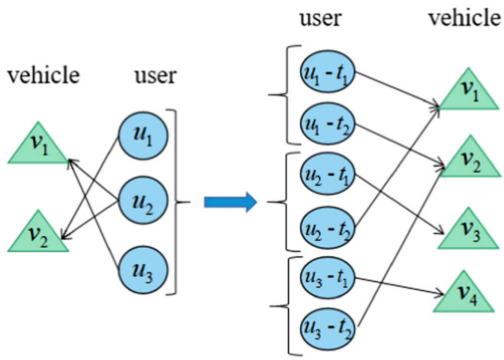


Fig. 5. Improvement of data structure.

Integration of the start and end positions of a user using a certain mode of travel: $\{v_i : l_j - > l_k\}$, The user departs at a certain time and arrives at a certain destination: $\{u_i : l_j - > l_k\}$. Concrete portrayal of the trajectory information of the user using the mode of travel and the construction of a heterogeneous information network.

4.2 Experimental Setup

Baselines. To verify the validity of our proposed methods, three different representatives recommended methods are considered in this paper. The first is based on a collaborative filtering approach: item KNN, MF, SVD. Then there are the more advanced neural network-based models: GMF, MLP, NeuMF. Finally, the HIN-based model: metapath2vec. As well as several variants of model DTN-MMTMRec, the following are specific descriptions of the various methods.
Classical recommendation models based on collaborative filtering:

- itemKNN [28]: This is a classic collaborative filtering approach that evaluates a user's preferences based on their history of interaction with an item, and then makes recommendations for them.
- MF [12, 29]: This is a standard matrix factorization method that projects users and items into the same latent space and uses latent feature vectors to represent them. Thereafter, the user's interaction with the item is modeled as the inner product of its latent vector.
- SVD [30]: This is a feature-based matrix decomposition method, where the object relationships of the heterogeneous information network in this paper are used as feature inputs.

Neural network-based collaborative filtering models:

- GMF: An example of neural collaborative filtering, applying a linear kernel to model latent feature interactions.
- MLP: An example of neural collaborative filtering, using a non-linear kernel to learn interaction functions from data.
- NeuMF [7]: This is a state-of-the-art neural network recommendation method that fuses GMF and MLP to better simulate complex user-item interactions.

HIN-based model and variants of model DTN-MMTMRec:

- metapath2vec [25]: The skip-gram model is used to update the node embeddings generated by the meta-path-guided random walks. The same MLP as in this paper is used here to predict the results.
- DTN–MMTMRec$_{mean}$: It is a variant of DTN-MMTMRec that uses the average aggregation function (Eq. 4) in both major processes of aggregation in the model to aggregate meta-path-guided neighbors to enrich the feature embedding of users and travel modes.
- DTN–MMTMRec$_{cnn}$: It is also a variant of DTN-MMTMRec, where we apply the convolution operation (Eq. 6) to aggregate embeddings between meta-paths.

- DTN–MMTMRec$_{lstm}$: It is a variant of DTN-MMTMRec, and in this model, we use LSTM (Eq. 5) to aggregate embeddings between meta-paths.
- DTN-MMTMRec: This is our complete model.

Evaluation Indicators. In this experiment, to evaluate the recommendation performance, we randomly divide the entire user travel trajectory data into a training set (80%) and a test set (20%), using the K-th precision (Prec@K) and the K-th recall (Recall). @K) and K-th normalized discounted cumulative gain (NDCG@K) are used as evaluation indicators. The larger the NDCG value, the better the performance. For stability, we use different random split training/test sets to run multiple times and average the results.

Implementation Details. We use the python library of Keras to implement the DTN-MMTMRec model. Firstly, the nodes in the heterogeneous information network of this article are initially embedded, and the embedding size is set to 64 dimensions, and then the users, as well as the travel mode sequences, are dynamically modeled and the different effects produced are analyzed by aggregating meta-path-guided neighborhood features using different aggregation functions. In the training phase, the model parameters were randomly initialized with a Gaussian distribution and the model was optimized using a small batch Adam [24], setting the batch size to 256 and setting the learning rate to 0.001. Besides, the number of sampled meta-path instances was four (as shown in Table 2). The other comparison methods followed the appropriate configuration and architecture accordingly and set the same evaluation metrics to facilitate the comparison of effectiveness.

4.3 Result Analysis

Table 3 shows the experimental results of the model as well as the comparison methods on the dataset, with the following main findings.

Table 3. Experimental results for each type of recommendation method on the dataset. The models with an '*' in the data are those with the best recommendation performance in each category, and we use bold to indicate the experimental results of the models. By comparing the metrics, the recommendation performance of the model proposed in this paper is more significant.

Method category	Model	Evaluating indicator		
		Precision@5	Recall@5	NDCG@5
CF	itemKNN	0.2372	0.3513	0.4821
	MF	0.2594	0.3750	0.5012
	SVD	0.2628	0.3757	0.5095*
NCF	GMF	0.2798	0.3951	0.5253
	MLP	0.2810	0.4030	0.5326
	NeuMF	0.2970	0.4127	0.5655*
HIN	Metapath2vec	0.3102	0.4134	0.5701
	DTN-MMTMRec	**0.3189**	**0.4279**	**0.6027**

According to the results of various experimental metrics, the a-model proposed in this paper is very effective for the multimodal travel mode recommendation task, and also demonstrates that the meta-path-guided graph neural network can well enrich the feature embedding of users and travel modes and improve the overall recommendation performance.

Based on the experimental results it can be observed that the meta-path-based model proposed in this paper outperforms the other recommendation methods compared. three recommendation methods based on collaborative filtering, where item KNN has the weakest performance. MF, SVD performance is not very different, which suggests that it is not sufficient to assess user preferences for travel modes based on the interaction between users, travel modes alone. The three examples of neural collaborative filtering frameworks in which a has performed well are based on the idea of modeling complex interactions between users and modes of travel using multi-layer perceptrons by replacing the inner product with a neural structure that can learn arbitrary functions from the data, with experimental results demonstrating the superiority of deep neural networks played in this regard. The DTN-MMTMRec model shows the best recommendation performance, indicating that it is feasible to apply the user's Spatial-temporal travel trajectory to a heterogeneous information network, and the meta-path-guided graph neural network is used to capture heterogeneous information, which produces a good recommendation effect.

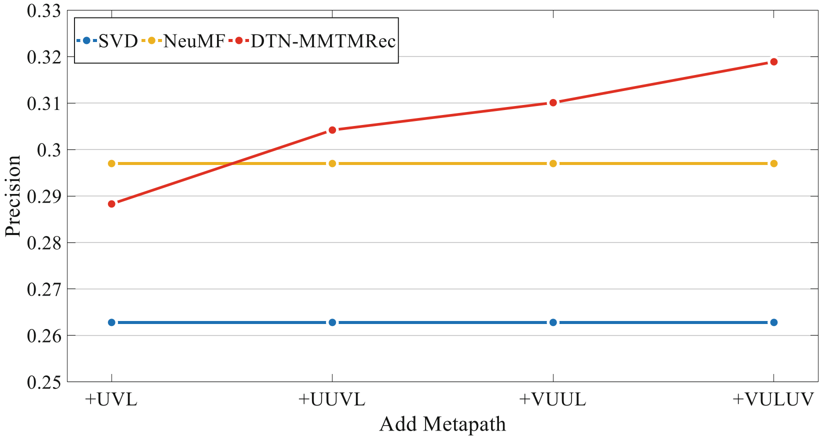
The model metapath2vec, also based on a heterogeneous information network and with the same meta-path set up in the experiment, has a lower recommended performance compared to our model. The DTN-MMTMRec model performs aggregation within meta-paths as well as aggregation between meta-paths along existing meta-paths, whereas metapath2vec uses a random wandering strategy to embed node features without learning semantic information between different meta-paths, so its recommendation performance is poor, and it also proves that our model has better recommendation results.

Table 4. Data on evaluation indicators for several variants of the DTN-MMTMRec model using different aggregation functions.

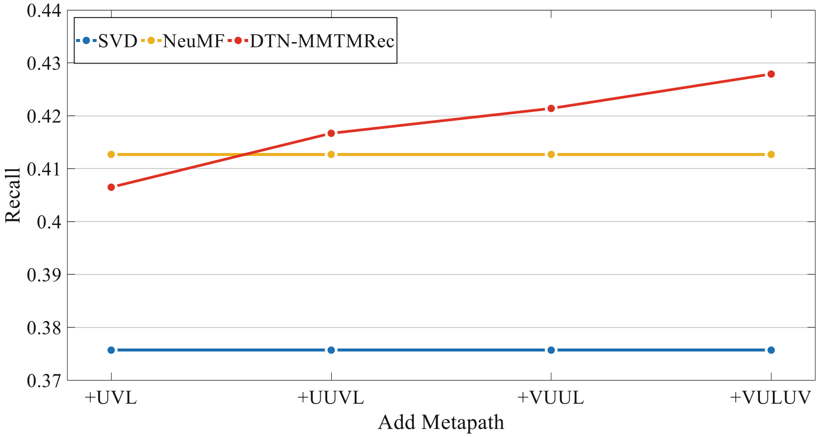
Model	Evaluating indicator		
	Precision@5	Recall@5	NDCG@5
$\mathrm{DTN-MMTMRec_{mean}}$	0.3065	0.4051	0.5611
$\mathrm{DTN-MMTMRec_{cnn}}$	0.3112	0.4125	0.5893
$\mathrm{DTN-MMTMRec_{lstm}}$	0.3162	0.4190	0.5972
$\mathrm{DTN-MMTMRec}$	**0.3189**	**0.4279**	**0.6027**

4.4 Result Analysis

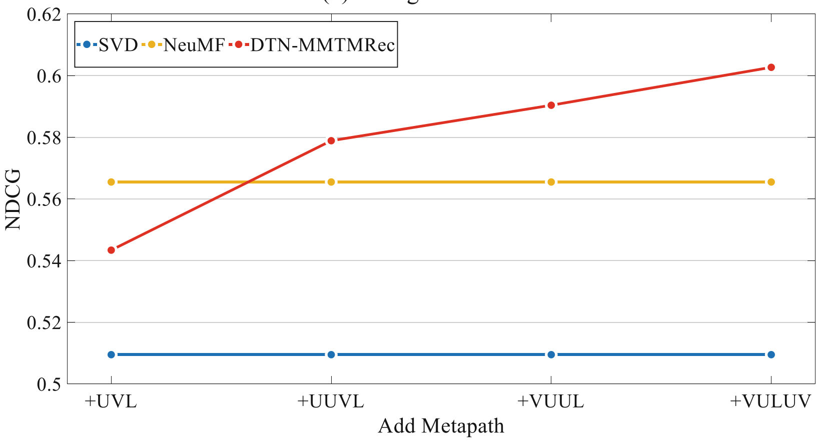
We select four meta paths with different semantics based on the relationships between objects in heterogeneous information networks and then use the aggregate meta paths of graph-neural networks to guide neighbors to improve recommendation performance. Because different meta paths have different semantics, we analyze the impact of different meta paths on the recommended performance by gradually incorporating existing meta paths into the DTN-MMTMRec model. The meta paths in this article are UVL, UVUV, VUUL, and VULUV, which are added to the model sequentially.



(a) Changes in Precision



(b) Changes in Recall



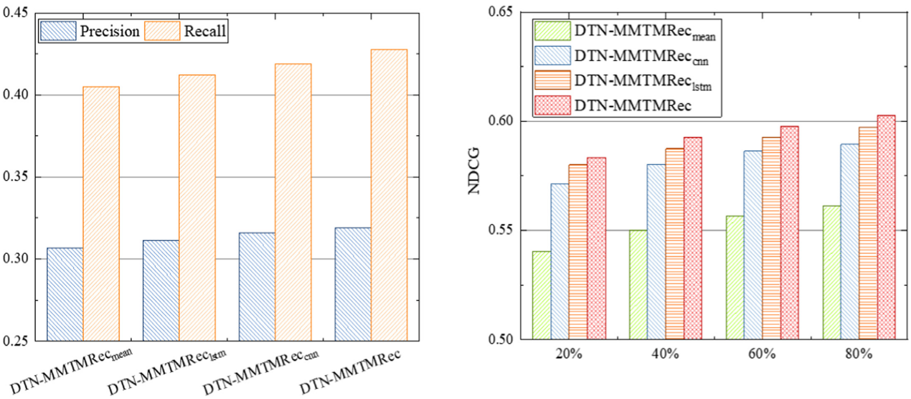
(c) Changes in NDCG

Fig. 6. Changes in each metric and comparison when adding meta-paths incrementally.

Figure 6 illustrates the change in performance of each metric of DTN-MMTMRec during the process of adding meta-paths. It can be observed that the recommendation performance of DTN-MMTMRec is significantly improved by adding more meta-paths. Also, the impact of different meta-paths on recommendation performance varies. In particular, we can notice a significant improvement in the performance of DTN-MMTMRec when adding UVUV, where similar users of the user are added, no longer just preferences embodied in the user's data, such as their nearby neighbors or friends, who are likely to have similar preferences to their mode of travel, which is where the semantic information extracted by the meta-path comes into play. The results show that adding meta-paths plays an important role in learning task-related embeddings and also illustrate that the heterogeneous graph neural network approach using meta-path guidance proposed in this paper is very effective for multimodal travel mode recommendations.

4.5 Effect of Aggregation Functions on Recommended Performance

The significance of meta paths is to extract specific interactions between different objects. The neighborhood information guided by aggregated meta-paths plays a key role, and we use different combinations of aggregation functions for comparison with the model constructed in this paper. The recommended performance can be found as DTN-MMTMRec > $\mathrm{DTN} - \mathrm{MMTMRec}_{\mathrm{lstm}}$>$\mathrm{DTN} - \mathrm{MMTMRec}_{\mathrm{cnn}}$>$\mathrm{DTN} - \mathrm{MMT}MRec_{\mathrm{mean}}$. $\mathrm{DTN} - \mathrm{MMTMRec}_{\mathrm{mean}}$ models use averaging functions at both stages of the aggregation process, ignoring the different semantics expressed by different meta-paths. The $\mathrm{DTN} - \mathrm{MMTMRec}_{\mathrm{cnn}}$ model uses convolutional features in the second stage of the aggregation process for both users and travel modes. The $\mathrm{DTN} - \mathrm{MMTMRec}_{\mathrm{lstm}}$ models all use LSTM dynamic modeling in the second stage of the aggregation process, while this paper's model uses LSTM to model users with temporal order and uses convolutional features to calculate the embedding of travel patterns, resulting in a significant improvement in recommendation performance.



(a) (b)

Fig. 7. Impact of different aggregation functions on recommendation performance.

To observe more clearly the effect of aggregation function on the performance of recommendations, this paper will illustrate it from two perspectives. Under the same conditions as the training set and the test set, the precision and recall performance indicators are compared using different aggregation function models (see Fig. 7(a)). We divide the data into five equal parts, one of which is used as a test set, the remaining data is stacked as a training set according to 20%, 40%, 60%, 80% and compared with different aggregators (see Fig. 7(b)). Here we can see that with the increase of training set data, the improvement of recommended performance decreases gradually, and the impact of different aggregators is also compared with Table 4. Echo.The results show that using heterogeneous information networks to obtain richer semantic information is very effective to improve the recommendation performance. The DTN-MMTMRec model proposed in this paper takes full advantage of the meta-path in heterogeneous information networks.

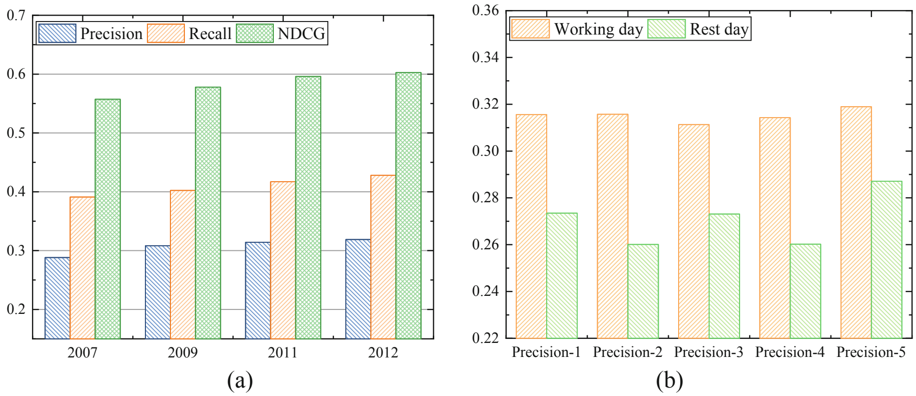


Fig. 8. Variation in performance metrics for model DTN-MMTMRec in specific data.

4.6 Performance of the Model on Specific Datasets

Over time, urban population density continues to increase and residents' travel needs are diversified. Our model has played an important role in the complex transportation network. Figure 8 shows the recommended effects of DTN-MMTMRec on different specific datasets. Figure 8 (a) shows that the recommended performance of the model has been improving over the data change from 2007 to 2012, which also proves that our model is more effective as travel demand increases. Figure 8 (b) compares the recommendation effect of DTN-MMTMRec on five different working days and rest days. It is not difficult to see that the recommendation performance of rest days is slower and less stable than that of working days, because the residents' travel on working days is more stable, and the travel demand on rest days varies greatly. The model DTN-MMTMRec is feasible in the recommendation of multimodal travel modes, can meet the personalized needs of daily travel of residents, and is more effective on weekdays.

5 Conclusion

In this paper, our proposed multimodal travel mode recommendation model based on dynamic traffic networks is feasible and effective. The model constructs the user's historical Spatial-temporal travel trajectory into a heterogeneous information network and provides insight into the user's preference for travel modes. Meta-path-guided graph neural networks for dynamic modeling of recommended objects play a key role in the overall recommendation process. At the same time, the real-world dataset is used and the data set is cleaned, filtered, and integrated with new data structure, and experimental evaluation shows that the proposed method in this paper produces good recommendation performance when applied to the recommendation of multimodal travel modes. However there is still relatively little work in this area, the aspects considered are not comprehensive enough and there is much room for improvement. On the one hand, the current work does not take into account users' interest in points of interest in their behavioral trajectories, the fact that different points of interest are related to travel modes, as well as the economic gap between users themselves and the impact of travel costs on travel mode choice. On the other hand, we can continue to improve the model framework, add attention mechanism, and select more important meta-paths to further improve the overall quality of travel mode recommendation.

Acknowledgment. This research was supported by research grants from the Beijing Natural Science Foundation (No. 4192004), the Project of Beijing Municipal Education Commission (No. KM201810005023, KM201810005024).

References

1. Enjian, Y., Weidi, C., Tianwei, L., Yang, Y.: Transportation mode selection model considering traveler's personal preferences. J. Beijing Jiaotong Univ. **209**(01), 46–52 (2020)
2. Kumar, G., Jerbi, H., Mahony, M.: Personalised recommendations for modes of transport: a sequence-based approach. In: International Workshop on Urban Computing at ACM SIGKDD. ACM (2016)
3. Bing-Feng, S., et al.: Urban multimodal traffic assignment model based on travel demand. China J. Highw. Transport **23**(6), 85–91 (2010)
4. Shi, C., et al.: A survey of heterogeneous information network analysis. IEEE Trans. Knowl. Data Eng. **29**(1), 17–37 (2016)
5. Sun, Y., et al.: RankClus: integrating clustering with ranking for heterogeneous information network analysis. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, Saint Petersburg, Russia (2009)
6. Zhi-Yan, F., et al.: Influence model of social network traffic information on the travel mode choice behavior. J. Transp. Syst. Eng. Inf. Technol. (2019)
7. Xiangnan, H., Lizi, L., Hanwang, Z., Liqiang, N., Xia, H.,Tat-Seng, C.: Neural collaborative filtering. In: 26th International Proceedings on World Wide Web, pp. 173–182 (2017)
8. Xiang, W., et al.: Neural Graph Collaborative Filtering, pp. 165–174 (2019)
9. Shi, C., et al.: Heterogeneous information network embedding for recommendation. In: IEEE Transactions on Knowledge and Data Engineering, pp. 1–1 (2017)

10. Fan, S., et al.: Metapath-guided heterogeneous graph neural network for intent recommendation. In: 25th ACM SIGKDD International Conference. ACM (2019)
11. Zhang, C., et al.: Heterogeneous graph neural network. In: 25th ACM SIGKDD International Conference. ACM (2019)
12. Xiangnan, H., HanWang, Z., Min-Yen, K., Tat-Seng, C.: Fast matrix factorization for online recommendation with implicit feedback. In: 39th International ACM SIGIR conference. ACM (2016)
13. Luan, K., Zhicai, J., Fang, Z.: Research on commuter's choice behavior between travel mode and trip chain. J. Highw. Transp. Res. Dev. (2010)
14. Kipf, T.N., Welling, M.: Semi-Supervised Classification with Graph Convolutional Networks (2016)
15. Binbin, H., et al.: Leveraging meta-path based context for top- N recommendation with a neural co-attention model. In: 24th ACM SIGKDD International Conference. ACM (2018)
16. Shi, C., et al.: Semantic path based personalized recommendation on weighted heterogeneous information networks. In: ACM International on Conference on Information and Knowledge Management. ACM (2015)
17. Liu, Z., et al.: Heterogeneous graph neural networks for malicious account detection. In: ACM International Conference, pp. 2077–2085. ACM (2018)
18. Zheng, Y., et al.: Learning transportation mode from raw GPS data for geographic applications on the web. In: 17th International Conference on World Wide Web, WWW '08, New York, USA, 2008, pp. 247–256 (2008)
19. Zheng, Y., Yukun, C., Quannan, L., Xing, X., Wei-Ying, M.: Understanding transportation modes based on GPS data for web applications. ACM Transactions on the Web (2010)
20. Zheng, Y., Quannan, L., Yukun, C., Xing, X., Wei-Ying, M.: Understanding mobility based on GPS data. In: 10th International Conference on Ubiquitous Computing, UbiComp'08, New York, USA, 2008, pp. 312–321 (2008)
21. Dabiri, S., Heaslip, K.: Inferring transportation modes from GPS trajectories using a convolutional neural network. Transp. Res. Part C Emerg. Technol. **86**, 360–371 (2018)
22. Zheng, Y., Xing, X., Wei-Ying, M.: Geolife: a collaborative social networking service among user, location and trajectory. In: IEEE Database Engineering Bulletin (2010)
23. Yu, Z., Lizhu, Z., Xing, X., Wei-Ying, M.: Mining interesting locations and travel sequences from GPS trajectories. In: 18th International Conference on World Wide Web, WWW '09, New York, USA, 2009, pp. 791–800 (2009)
24. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. Comput. Sci. (2014)
25. Dong, Y., Chawla, N.V., Swami, A.: Metapath2vec: Scalable Representation Learning for Heterogeneous Networks. ACM (2017)
26. Xiaohua, Y., et al.: A study of multimodal traffic assignment based on a multi-level network. J. Transp. Inf. Saf. **36**(001), 103–110 (2018)
27. Hamilton, W.L., Ying, R., Leskovec, J.: Inductive Representation Learning on Large Graphs (2017)
28. Sarwar, B., et al.: Item-based Collaborative Filtering Recommendation Algorithms. ACM (2001)
29. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. Computer **42**(8), 30–37 (2009)
30. Chen, T., et al.: SVDFeature: a toolkit for feature-based collaborative filtering. J. Mach. Learn. Res. **13**(1), 3619–3622 (2012)