Fusing User Preferences and Spatiotemporal Information for Sequential Recommendation

SIZHE YIN, YANG XIA, YUJING LIU, SONGHE HAN, AND ZIJIE OUYANG
School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China
Mine Digitization Engineering Research Center, Ministry of Education of China, China University of Mining and Technology, Xuzhou 221116, China
Corresponding author: Yang Xia (yxia@cumt.edu.cn)
This work was supported in part by the National Natural Science Foundation of China under Grant 51874300.

ABSTRACT At present, the research on sequence recommendation mainly focuses on using the historical interaction data between users and items to mine their relationship, so as to predict the next interaction between users and items, then generate the personalized recommendation. Spatiotemporal information is very important to further improve the accuracy and quality of recommendation, but the existing sequence recommendation models are mainly based on recurrent neural network (RNN), and pay less attention to spatiotemporal information. Most of the recommendation models are still in the early stage of merging spatiotemporal context information, and the processing effect of long sequence data is not ideal. A sequential recommendation model integrating user preferences and spatiotemporal information is proposed. The model captures user item long-term preferences through spatiotemporal GRU algorithm and user item short-term preferences through attention mechanism. Finally, the learned long-term and short-term preference features and user portrait features are combined to predict the next recommendation location. The experimental results on two real data sets Foursquare and Brightkite show that the proposed model performs better than state-of-the-arts in three evaluation indicators HR@K, NDCG@K and MAP@K.

INDEX TERMS Recommendation algorithm, preference learning, spatiotemporal information, sequential recommendation.

I. INTRODUCTION
With the rapid development of Internet information technology in recent years, people’s clothing, food, housing and transportation are also closely related to the Internet, which brings the problem of data overload. How to help users accurately search for items of interest in massive data has become an important topic, and the emergence of recommendation system can well alleviate this problem. At present, recommendation system has been widely used in e-commerce, news, online advertising, music, movies, social networks and other Internet applications.

The research on recommendation system can generally be divided into collaborative filtering (CF) [1], content-based [2], and hybrid methods [3]. Over the years, people have proposed many powerful neural network recommendation algorithms [4], [5], [6]. These recommendation algorithms improve the accuracy, but there are also some deficiencies in dynamic recommendation. The item recommended by the user-based CF can be explained as “the users similar to you like this item”. The item-based CF can be explained as “this item is similar to the item which you previously liked”. Although CF ideas have significantly improved in accuracy, they are not as intuitive as content-based algorithms [7]. CF achieved further success after integration with latent factors model (LFM). In many LFM, matrix factorization (MF) [8] and its variants are particularly successful in evaluating prediction tasks. However, the potential factors in LFM have no intuitive meaning, which makes it difficult to understand why a item has a good prediction or why it is recommended. Sequential recommendation aims to explicitly model the sequence behavior of users. It can be defined as learning the time dynamics of user behavior in sequence data and predicting the items that users want to click later, that is,
user behavior. In most cases, there is a time order relationship. What items are recommended to users at a certain time is generally decided according to the user's behavior before the current time. User-item interactions are inherently sequential. In the real world, users' shopping behavior usually occurs continuously in order rather than in an isolated way. Secondly, the preferences of users and the popularity of goods change over time, rather than static. In fact, users’ preferences and tastes may change over time. Assuming that a user behavior sequence is [T-shirts, pants, sneakers, apple phones, headphones, laptops], it can be seen that users pay more attention to clothes in the previous period and are more interested in electronic products in the latter period. Therefore, it is necessary to do corresponding processing according to different time periods of user behavior in order to significantly improve the performance of the recommendation system.

Reference [9] summarizes the sequence recommendation system, including traditional sequence models, such as sequence mining pattern and Markov chain model. Potential representation models, such as factorization machines, are embedded. Deep neural network models such as CNN, RNN, GRU. And advanced attention networks, memory networks and hybrid models. It can be seen that there are many methods in the sequence recommendation system, and each method has its own characteristics. Many methods mine user behavior sequences to better understand users/items and generate the next recommendation. However, in the existing sequence recommendation system, the use of spatio-temporal information is insufficient, the interaction of mining user item features is not comprehensive, and the effect of processing long sequence data is not ideal. In addition, the generated next recommendation is not personalized enough and the amount of information is not enough to arouse the interest of users. The quality and accuracy of recommendation need to be improved. This paper mainly optimizes the accuracy and quality of recommendation.

In order to solve the above problems, we design a novel sequential recommendation model called User Preferences and Spatiotemporal Information for Sequential Recommendation (UTSR), which improves the recommendation quality and accuracy of the recommendation system by fusing user preferences and spatiotemporal information. Our main contributions are summarized as follows:

- We analyze and evaluate the experimental results. The results show that compared with the baseline model, our method improves the quality and accuracy of recommendation, so as to generate personalized, rich and high-quality next recommendation.

The rest of this paper is organized as follows. Section 2 introduces the related work. Section 3 introduces the problem expression and the proposed UTSR model. Section 4 introduces the data sets and experimental methods used in this paper. A large number of experimental results are discussed. Finally, we summarize our work in Section 5.

II. RELATED WORK

A. SEQUENCE RECOMMENDATION SYSTEMS

The existing sequence recommendation systems mainly focus on predicting the identity of the next item that users may interact with. The noteworthy work includes sequence pattern mining and Markov chain (MC), which aims to capture the high-order relationship between users and items. Wang et al. [10] proposed an improved model HRM of FPMC model, which essentially adds nonlinear transformation to FPMC. In recent years, sequence recommendation model based on deep neural network has become popular. GRU4Rec [11] used the cyclic neural network in sequence recommendation for the first time, modeled all sessions, solved the problem of modeling sparse sequence data, and designed the ranking loss function suitable for the recommendation task. Tang and Wang [12] proposed a model called Caser, which extracts the information of short-term sequence through convolution neural network (CNN), constructs the vector of historical behavior into a matrix, then takes this matrix as an image in time and potential space, and finally obtains the user’s short-term representation through convolution operation. The SR-GNN model proposed by Wu et al. [13] models the session sequence as graph structure data. Based on the session graph, GNN can capture the complex transformation of items. Each session uses the attention mechanism to combine the overall preference with the current preference. With the emergence of Transformer [14], the self-attention mechanism has become the mainstream modeling method of sequence recommendation. The SASRec model proposed by Kang and McAuley [15] introduces transformer into sequence recommendation. SASRec is essentially a non-personalized model because it does not include the embedding of personalized users. In order to solve this problem, Wu et al. Proposed a personalized transformer model SSE-PT [16], which improves the performance by 5% compared with SASRec under the same evaluation index. Experiments show that SSE-PT is more explanatory, and each recommendation for users will focus on the user’s recent behavior pattern. The AttRec model proposed by Zhang et al. [17] models users’ short-term preferences through self-attention mechanism and users collaborative metric learning to model users’ long-term preferences. The BERT4Rec model proposed by Sun [18] et al. applies BERT to sequence recommendation for the first time, and introduces the close task to replace the goal.
of a single task to solve the information leakage that may be caused by training in the two-way depth model. Lin et al. [19] proposed the fusion item similarity model and self attention network model (FISSA). The model designs a global representation learning module to effectively capture users’ global preferences. It can be regarded as a location-based attention layer, which is well consistent with the parallel training process of self attention framework. It also designs a gating module based on GRU, which balances the local and global representation by considering the information of candidates, so as to deal with the uncertainty of user intention at the same time. The multi interest dynamic routing recommendation model (MIND) proposed by Li et al. [20] designs a vector layer to extract users’ multi interests based on the capsule routing mechanism, which aggregates users’ history into some clusters. In this way, for a user, MIND can output multiple vectors to express the user’s multiple interests. The adjustable multi interest recommendation framework model (CoMiRec) proposed by Cen et al. [21] captures multiple interests from the user behavior sequence, which can be used to retrieve candidate items from the large-scale item pool, and then input these items into the aggregation module to obtain the overall recommendation. The aggregation module uses controllable factors to balance the accuracy and diversity of recommendations. There are some defects in some existing methods. For example, not all historical behaviors have an effect on the user’s future preferences, and the user’s historical behavior is represented by an embedding, ignoring the relationship between various items. To solve this problem, Chen et al. [22] proposed a memory network model (MANN), which introduces the memory network into the sequence recommendation for the first time, and directly captures the dependency between any historical user item interaction by merging the external memory matrix. Such a matrix can store and update the historical interaction in a sequence more clearly and dynamically, so as to improve the expression ability of the model and reduce the interference of irrelevant interaction. Tang et al. [23] proposed a neural hybrid recommendation algorithm (M3) for long-distance correlated user sequences. The algorithm adopts hybrid models, and each model has a different time range. These models are combined by a learning gating mechanism, which can use different model combinations for sequence recommendation given different context information. Enhance the quality of our previous work based on the integration of temporal and spatial information and user preferences.

B. SPATIOTEMPORAL INFORMATION ANALYSIS

In the sequence recommendation system, it is very important to mine the order dependence between different behaviors. One of the important methods is to consider spatiotemporal information. Li et al. [24] proposed a self attention sequence recommendation (TiSASRec) model based on time interval perception. The model models the timestamp in interaction into the sequence model framework to explore the impact of different time intervals on predicting the next item. Wu et al. [25] proposed the framework of situational time attention mechanism (CTA). The architecture learns to weigh the impact of historical behavior, including not only what behavior it is, but also when and how the behavior occurs. It also inherits the advantages of self attention mechanism, reduces parameters and improves computational efficiency, because the architecture can also be deployed in parallel and interpretable. The attention mechanism proposed by Cho et al. [26] and the mixed (MEANTime) model of multi time embedding. The model adopts various types of time embedding and operates multiple self attention heads at the same time. Each head uses position embedding to extract specific patterns from the user’s behavior.

Wu et al. [27] proposed personalized long-term and short-term preference learning model (PLSPL). In the short-term module, in order to better understand the different influences of locations and categories of POIs, two LSTM models are trained for location and category-based sequence respectively. Zhao et al. [28] proposed a new Spatio-Temporal Gated Network (STGN) by enhancing long-short term memory network, where spatio-temporal gates are introduced to capture the spatio-temporal relationships between successive check-ins. In order to alleviate the low performance of model training caused by the sparsity of user check-in data, Liu et al. [29] proposed a class aware gated recursive unit (CA-GRU) model, which captures the long-range dependence between user check-ins and get better recommendation results of POI category. Huang et al. [30] proposed an attention-based spatiotemporal LSTM (ATST-LSTM) network for next POI recommendation. ATST-LSTM can focus on the relevant historical check-in records in a check-in sequence selectively using the spatiotemporal contextual information. Wang et al. [31] proposed an Attention Sequential model based on Graph Neural Network (ASGNN) for accurate next POI recommendation. ASGNN models user’s check-in sequences as graphs and then use Graph Neural Networks (GNN) to learn the informative low-dimension latent feature vectors of POIs. Chen et al. [32] proposed a RNN-based next POI recommendation approach that considers both the location interests of similar users and contextual information. Sun et al. [33] proposed a novel method named Long-and Short-Term Preference Modeling (LSTPM). LSTPM consists of a nonlocal network for long-term preference modeling and a geo-dilated RNN for short-term preference learning. The overall structure of this model is relatively simple.

In this paper, we implement UTSR based on the improved GRU formula. Due to the particularity of sequence recommendation task, we need to mine the correlation between the data before and after the sequence. The common method to process sequence information is to use recurrent neural network RNN, but it lacks the ability to filter data and has some problems such as gradient explosion. The emergence of LSTM effectively alleviates these problems. Due to its own gating device, LSTM will selectively store information. GRU, as a variant of both, can achieve considerable results,
and is easier to train, which can greatly improve the training efficiency. In recent years, the recommendation model based on GRU and its variants has achieved the best results in a large number of recommendation tasks.

III. UTSR MODEL

Our goal is to establish a sequential recommendation system model, which uses user preferences and spatiotemporal information to improve the accuracy and quality of recommendation. When recommending the next location to the user, the system will analyze the previous user item interaction records to extract the user’s long-term preferences, and then combine the continuous time factors to generate the location that the user may be most interested in in the near future.

A. SPATIOTEMPORAL INFORMATION DEFINITION

Firstly, the influence of time information is mainly reflected in four aspects: absolute time factor, continuous time factor, time similarity relationship and periodic time mode. Secondly, we define the spatial factor as the distance factor. Because GRU cannot model in continuous time, this paper discretizes time and uses time interval as the time factor.

In this paper, we define a specific time transformation matrix $T_{t-i}$ for the time interval $t - i$ before the current time $t$, and a specific distance transformation matrix $S_{U_l U_l}$ according to the two geographic distances under the coordinates. The matrix $T_{t-i}$ captures the impact of the most recent elements in the history, also taking into account specific time intervals. $q_t$ represents the coordinates of the location the user $U$ is visiting at the time $t$, and the geographic distance is calculated by the Euclidean formula, as shown in equation (1):

$$U_l - U_l = \|x_l^u - x_l^u, y_l^u - y_l^u\|_2$$

(1)

Then the new candidate state vector in time $t$ is obtained as follows in equation (2):

$$\hat{C}_t = tanh(W_{c1}U_l + W_{c2}S_{U_l U_l} + W_{c3}T_{t-i} + W_{c4} \odot h_{t-1} + b_c)$$

(2)

In addition, we define a user set $U$ and an item set $I$, $X^u = (X_{u1}, X_{u2}, \ldots, X_{uO})$ represents the sequence of items that the user has interacted with before, $X_{ui} \in I$, and the following table index of $X_{ui}$ represents the order in which the item interactions appear in the sequence. Figure 1 shows the structure of a sequence recommendation model UTSR that fuses user preferences and spatiotemporal information. The model is divided into an embedding layer, an interest mining layer, a feature vector fusion layer, and a fully connected layer. The interest mining layer consists of attention mechanism and improved GRU algorithm. The attention mechanism is used to capture the user’s short-term interest, and the improved GRU algorithm is used to capture the user’s long-term interest. And the item embedding vectors of the two inputs are shared.

B. EMBEDDING LAYER

The input of the input layer is preprocessed to generate user feature items and user-item interaction feature items. The method is to embed the high-dimensional sparse one-hot code into a low-dimensional dense feature vector and then input it into the model. This processing method can greatly reduce the amount of calculation and can also generate semantic correlation between the feature vectors. The user embedding matrix is $E_u$, and the user behavior sequence embedding matrix is $E_x \in \mathbb{R}^{n \times k}$, $e_u$ and $e_x$ represent the embedding vector representation of the user and the user behavior sequence, respectively.

C. INTEREST MINING LAYER

According to the above interpretation of spatiotemporal information, an improved GRU algorithm can be obtained, as shown in equations (3) to (6):

$$i_t = \sigma(W_{i1}U_l + W_{i2}h_{t-1} + W_{i3}S_{U_l U_l} + W_{i4}T_{t-i} + b_i)$$

(3)

$$f_t = \sigma(W_{f1}U_l + W_{f2}h_{t-1} + W_{f3}S_{U_l U_l} + W_{f4}T_{t-i} + b_f)$$

(4)

$$\hat{C}_t = \tanh(W_{c1}U_l + W_{c2}S_{U_l U_l} + W_{c3}T_{t-i} + W_{c4} \odot h_{t-1} + b_c)$$

(5)

$$C_t = i_t \odot \hat{C}_t + (1 - i_t) \odot C_{t-1}$$

(6)

where $\sigma$ is the sigmoid function, $W_{i1}, W_{i2}, W_{i3}, W_{i4}, W_{f1}, W_{f2}, W_{f3}, W_{f4}, W_{c1}, W_{c2}, W_{c3}, W_{c4}$ are transformation matrices, $b_i, b_f, b_c$ are bias vectors, $W_{i1}, W_{i4}, W_{f4}, W_{f1}, W_{f3}, W_{c4}$ are transition matrices of $S_{U_l U_l}$ and $T_{t-i}$. Therefore, $C_t$ includes not only the information of the original input $U_l$, but also the important information of the distance context $S_{U_l U_l}$ and the time context $T_{t-i}$. This enables the user’s preference in each hidden state to be enhanced. Finally, in a long sequence, the output at time $t$ is expressed as in equation (7):

$$h_t = \tanh(C_t)$$

(7)

In addition, we believe that the basic encoder-decoder has limitations. The biggest limitation is that the only connection between encoding and decoding is a fixed-length semantic vector. That is, the encoder compresses the entire sequence of information into a fixed-length vector. However, there are two drawbacks in this way. One is that the semantic vector cannot fully represent the information of the entire sequence, and the other is that the information carried by the first input content will be diluted, or in other words, covered by the later input information. The longer the input sequence, the more severe this phenomenon is. This makes it impossible to obtain enough information of the input sequence at the beginning of decoding, so the accuracy of decoding will naturally decrease. So an attention mechanism is introduced to capture sequential features from users’ short-term behavior history to fully understand and express their short-term interests. The calculation process of the self-attention mechanism is shown
where \( \tilde{e}_x \) represents the output result of the multi-head self-attention mechanism, \( N \) is the total number of head spaces and \( W^N \) is a parameter. Because the dimension size of the feature vectors of multiple head spaces after splicing is generally not equal to the embedding dimension of the original feature vector, the projection matrix is \( W^N \) is used to restore it to the dimension size of the original vector.

\[
G = \text{Concat}(\tilde{e}_x^N, h_t, e_u)
\] (16)

D. FEATURE VECTOR FUSION LAYER

The main task of this section is to perform multimodal fusion of the user’s long-term and short-term preferences and user portrait features obtained through the improved GRU algorithm and attention mechanism above. As shown in the equation (16):

\[
\text{Dice}(x) = p(x) + (1 - p(x))ax 
\] (17)

\[
p(x) = \frac{1}{1 + \exp\left(-\frac{x - E(x)}{\sqrt{var(x) + \epsilon}}\right)}
\] (18)

In this paper, we regard the recommendation prediction task as a binary classification task, and finally select the Softmax function to perform binary prediction on the last hidden layer. As shown in the equation (19):

\[
\hat{y} = \text{Softmax}(W_HD_H + b_H)
\] (19)

where \( W_H \) is the trainable parameter matrix, \( b_H \) is the bias, and \( D_H \) is the hidden unit of the output of the \( H \) layer. \( \hat{y} \in (0, 1) \) Represents the probability of recommending the next item. Since the recommendation task is regarded
as a binary classification task, we choose the commonly used cross-entropy loss function as the optimized objective function. As shown in the equation (20):

$$L = -\frac{1}{N} \sum_{i=1}^{N} (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i))$$ (20)

where $y_i \in (0, 1)$ indicates the real recommendation situation. If $y_i = 0$, it indicates that the item is not recommended, otherwise, it indicates that it is recommended. $N$ indicates the batch size for training.

### IV. EXPERIMENT

In this section, we conduct experiments to evaluate our proposed model. First, we introduce the dataset, baseline models for comparison, and evaluation metrics. Second, we present and analyze the experimental results through a series of evaluation metrics. Finally, the effectiveness of our method in generating the next recommendation is further analyzed through case studies.

#### A. DATA SET

In our experiments, we selected two widely used public datasets, Foursquare and Brightkite. The Foursquare dataset is a large-scale location-based social networking site that allows users to check in at different locations. It analyzes the spatial information, temporal relationship, social connection, text content and popularity information about the check-in data. In the Foursquare dataset, the time range is April 2013 to October 2014. The Brightkite dataset is user check-in information provided by a location-based social networking platform. Users share location information when checking in, so each check-in data has a location information.

In addition, to alleviate data sparsity and cold-start issues, users and points of interest with less than 5 check-in records are removed.

#### B. CONTRAST ALGORITHM

In this subsection, we introduce the baseline models used for comparison. To evaluate the accuracy of rating predictions, we compared our model with the following models.

1) FPMC
   Factorizing Personalized Markov Chains [35], this method is to capture the user’s long-term and short-term interest preferences by combining matrix factorization with Markov chains, and predict the next recommendation.

2) BPR
   Bayesian personalized ranking [36], a generalized criterion and learning algorithm for personalized ranking, which we use in the next recommendation.

3) RNN
   Recurrent Neural Networks [37], this method captures dynamic information in user behavior sequences through a standard recurrent structure, and can classify serialized data.

4) GRU
   Gate Recurrent Unit [38], GRU network is a more robust variant of RNN, which is more advantageous in capturing long-term dependencies.

5) ST-RNN
   Spatial Temporal Recurrent Neural Networks [39], this method extends RNN, which can model local temporal and spatial context in each layer.

6) SASRec
   Self-Attention Based Sequential model [15], this method introduces Transformer into sequence recommendation, and many models are currently improvements to SASRec.

7) FPMC-LR
   Factorizing Personalized Markov Chains And Localized Regions [40], this model is a new new matrix factorization method, which considers the movement of the user’s local area, and uses the information of the local area to reduce the computational cost.

8) PRME-G
   Personalized Ranking Metric Embedding-Geographical [41], this model is an advanced Markov chain method that incorporates sequence information, personal preferences, and geographic influences.

#### C. EVALUATION METRICS

In order to verify the effectiveness of the UTSR model, this paper adopts the Top-K evaluation index to measure the recommendation performance of the model, including Hit Rate (HR), Normalized Cumulative Gain (NDCG), and Average Accuracy Rate (MAP). HR@K measures whether a test item appears in the top K items of the predicted recommendation list. NDCG@K measures whether the test item appears in the top K items in the top ranking, focusing on sequentiality. MAP@K measures the average learning accuracy of all test categories on the recommendation model. As shown in equation (21) to (24):

$$HR@K = \frac{\sum \text{Hits@K}}{TN}$$ (21)

$$NDCG@K = Z_k \sum \frac{2^{r_i} - 1}{\log_2(i + 1)}$$ (22)
where $K$ represents the number of recommended items, $|TN|$ represents the number of test sets, and the numerator is the cumulative number of test sets that exist in each user’s previous $K$ item list. $r_i$ is represented as the correlation at location $i$, if the item at location $i$ is in the test set, $r_i = 1$, otherwise $r_i = 0$. $Z_u$ is the regularization coefficient. $\Omega_i$ indicates the result of ground-truth, which can be understood as the correct label or standard answer. $p_{uj}$ indicates the position of the item $j$ in the recommendation list, $h(p_{uj} < p_{ui})$ means that the item $j$ ranks before the item $i$ in the recommendation list. $MAP$ indicates the re-average of $AP$ of all users $U$. In this paper, we choose $K = \{10, 20\}$ to illustrate the effects of three evaluation metrics under different recommendation numbers.

**D. EXPERIMENTAL SETUP**

We randomly split each dataset into three subsets, 80%, 10%, and 10%, for training, validation, and testing. The validation set is used to tune hyperparameters. For all experimental methods, the learning rate is set to 0.01, the embedding dimension is set to 100, the batch size is set to 256, the regularization coefficient $L_2$ is selected between $[1 \times 10^{-6}, 1 \times 10^{-5}, 1 \times 10^{-3}]$, and the model is optimized using the Adam function. The model is implemented based on TensorFlow 2.0 and Python 3.6.

**E. EXPERIMENTAL RESULTS AND ANALYSIS**

In order to verify the effectiveness of the UTSR model, this paper will conduct experimental analysis on it from three aspects. First, the UTSR model proposed in this paper is evaluated on two datasets with three evaluation metrics and eight comparison methods. Second, an experimental comparison of UTSR and its two variants is performed. Finally, the effects of different embedding dimensions on model performance are analyzed.

1) **ANALYSIS OF THE RESULTS OF THE COMPARISON ALGORITHM**

Tables 2 and 3 show the experimental results of all methods under the two datasets. The UTSR model outperforms all other compared methods on both datasets, validating their implicit preferences. FPMC-LR outperforms the base FPMC model, but does not incorporate the consideration of geographic information. PRME-G is an embedding method that ignores the consideration of visualization of the relationship between users and items.

2) **THE COMPARISON OF VARIANTS OF UTSR MODEL**

In order to verify the impact of the improved GRU algorithm on the model, this paper first compares the effect of our improved algorithm with the ordinary GRU network on the basis of the UTSR model, and then conducts an experimental comparison of UTSR and its two variants. Namely UTSR-GRU (replace the model algorithm with a normal GRU network), UTSR-RNN (replace with a simple recurrent neural network), UTSR-LSTM (replace with a long short-term memory neural network). The model depth remains the same, and the comparison results are shown in Table 4.

As can be seen from Table 4, the improved algorithm is more effective than the ordinary GRU network on both
TABLE 4. Comparative experiments based on three evaluation indexes on two data sets (all values are in percentages).

| Dataset | Method    | HR@10 | NDCG@10 | MAP  |
|---------|-----------|-------|---------|------|
| Foursquare | UTSR-RNN  | 0.5179| 0.4297  | 0.2915|
|         | UTSR-LSTM| 0.5361| 0.4539  | 0.3021|
|         | UTSR-GRU | 0.5648| 0.4617  | 0.3143|
|         | UTSR     | 0.5835| 0.4891  | 0.3188|
| Brightkite | UTSR-RNN  | 0.3214| 0.2419  | 0.1472|
|          | UTSR-LSTM| 0.3367| 0.2516  | 0.1538|
|          | UTSR-GRU | 0.3455| 0.2761  | 0.1585|
|          | UTSR     | 0.3623| 0.2977  | 0.1621|

FIGURE 2. Impact of different embedding dimensions on HR on data set Foursquare.

FIGURE 3. Impact of different embedding dimensions on NDCG on data set Foursquare.

FIGURE 4. Impact of different embedding dimensions on HR on data set Brightkite.

FIGURE 5. Impact of different embedding dimensions on NDCG on data set Brightkite.

3) THE EFFECT OF EMBEDDING DIMENSION SIZE

We further investigate the effect of the embedding dimension size on the performance of the UTSR model. In general, more embedding dimensions can improve the performance of the model. However, it can also lead to overfitting. Here, we set the range of dimensionality to 10-150 and use HR@K and NDCG@K to calculate the generalization ability of the model.

for each case. Through Figure 2-Figure 5, we observe that our UTSR model achieves stable performance in the range of 90-150 and 110-150 on both datasets, respectively. Therefore, on the Foursquare and Brightkite datasets, we can set the embedding dimension to 90 and 110 in our experiments, respectively.

V. CONCLUSION

In this paper, a sequential recommendation model (UTSR) that integrates user preferences and spatiotemporal information is proposed, which extracts long-term behavioral sequence features through an improved GRU algorithm, and extracts short-term behavioral sequence features through an attention mechanism, while taking into account the user portrait features. The combination of the three can better solve the sequence recommendation problem. Experiments on two public datasets show that UTSR outperforms other methods and validates its effectiveness. In the future, we will consider adding item labels, social information, etc. to the recommendation model to further improve the recommendation performance of the model.

REFERENCES

[1] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry, “Using collaborative filtering to weave an information tapestry,” Commun. ACM, vol. 35, no. 12, pp. 61–70, 1992.

[2] M. J. Pazzani and D. Billsus, “Content-based recommendation systems,” in The Adaptive Web (Lecture Notes in Computer Science), vol. 4321. Berlin, Germany: Springer, 2007, pp. 325–341.
[3] R. Burke, “Hybrid recommender systems: Survey and experiments,” User Model. User-Adapted Interact., vol. 12, no. 4, pp. 331–370, Nov. 2002.

[4] X. He and T.-S. Chua, “Neural factorization machines for sparse predictive analytics,” in Proc. 40th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., Aug. 2017, pp. 355–364.

[5] Y. He, L. Liao, H. Zhang, L.Nie, X. Hu, and T.-S. Chua, “Neural collaborative filtering,” in Proc. 26th Int. Conf. World Wide Web, Apr. 2017, pp. 173–182.

[6] W. Ma, M. Zhang, C. Wang, C. Luo, Y. Liu, and S. Ma, “Your tweets reveal what you like: Introducing cross-media content information into multi-domain recommendation,” in Proc. 27th Int. Joint Conf. Artif. Intell., Jul. 2018, pp. 3484–3490.

[7] J. Li, H. Riehl, H. A. Konstan, and J. Riedl, “Explaining collaborative filtering recommendations,” in Proc. ACM Conf. Comput. Supported Cooperat. Work (CSCW), 2000, pp. 241–250.

[8] Y. Koren, R. Bell, and C. Volinsky, “Matrix factorization techniques for recommender systems,” IEEE Comput., vol. 42, no. 8, pp. 30–37, Aug. 2009.

[9] S. Wang, L. Hu, Y. Wang, L. Cao, Q. Z. Sheng, and M. Orgun, “Sequential recommender systems: Challenges, progress and prospects,” 2019, arXiv:2001.04830.

[10] P. Wang, J. Guo, Y. Lan, J. Xu, S. Wan, and X. Cheng, “Learning hierarchical representation model for nextbasket recommendation,” in Proc. 38th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., Aug. 2015, pp. 403–412.

[11] B. Hidasi, A. Karatzoglou, L. Baltrunas, and D. Tikk, “Session-based recommendations with recurrent neural networks,” in Proc. Comput. Conf., Aug. 2015, pp. 1–10.

[12] J. Tang and K. Wang, “Personalized top-N sequential recommendation via convolutional sequence embedding,” in Proc. 11th ACM Int. Conf. Web Search Data Mining, Marina Del Rey, CA, USA, Feb. 2018, pp. 565–573.

[13] S. Wu, Y. Tang, Y. Zhu, L. Wang, X. Xie, and T. Tan, “Session-based recommendation with graph neural networks,” in Proc. 33rd AAAI Conf. Artif. Intell., Jul. 2019, vol. 33, no. 1, pp. 346–353.

[14] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” 2017, arXiv:1706.03762.

[15] W.-C. Kang and J. McAuley, “Self-attentive sequential recommendation,” in Proc. IEEE Int. Conf. Data Mining (ICDM), Nov. 2018, pp. 197–206.

[16] L. Wu, S. Li, C. J. Hsieh, and J. Sharpnack, “SSE-PT: Sequential recommendation via personalized transformer,” in Proc. 14th ACM Conf. Recommeder Syst., Sep. 2020, pp. 328–337.

[17] S. Zhang, Y. Tay, L. Yao, A. Sun, and J. An, “Next item recommendation with self-attentive metric learning,” in Proc. 33th AAAI Conf. Artif. Intell., vol. 9, 2019.

[18] F. Sun, J. Liu, J. Wu, C. Pei, X. Lin, W. Ou, and P. Jiang, “BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer,” in Proc. 28th ACM Int. Conf. Inf. Knowl. Manage., Nov. 2019, pp. 1441–1450.

[19] J. Lin, W. Pan, and Z. Ming, “FISSA: Fusing item similarity models with self-attention networks for sequential recommendation,” in Proc. 14th ACM Conf. Recommeder Syst., Sep. 2020, pp. 130–139.

[20] C. Li, Z. Liu, M. Wu, Y. Xu, H. Zhao, P. Huang, G. Kang, Q. Chen, W. Li, and D. L. Lee, “Multi-interest network with dynamic routing for recommendation at tmall,” in Proc. 28th ACM Int. Conf. Inf. Knowl. Manage., Nov. 2019, pp. 2615–2623.

[21] J. Chen, J. Zhang, X. Zou, C. Zhou, H. Yang, and J. Tang, “Controllable multi-interest framework for recommendation,” in Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Aug. 2020, pp. 2942–2951.

[22] X. Chen, H. Xu, Y. Zhang, J. Tang, Y. Cao, Z. Qin, and H. Zha, “Sequential recommendation with user memory networks,” in Proc. 11th ACM Int. Conf. Web Search Data Mining, Feb. 2018, pp. 355–364.

[23] J. Tang, F. Belletti, S. Jain, M. Chen, A. Beutel, C. Xu, and E. H. Chi, “Towards neural mixture recommender for long range dependent user sequences,” in Proc. World Wide Web Conf. (WWW), 2019, pp. 1782–1793.

[24] J. Li, Y. Wang, and J. McAuley, “Time interval aware self-attention for sequential recommendation,” in Proc. 13th Int. Conf. Web Search Data Mining, Jan. 2020, pp. 322–330.

[25] J. Wu, R. Cai, and H. Wang, “Dějú vyu: A contextualized temporal attention mechanism for sequential recommendation,” in Proc. 29th Int. Conf. World Wide Web (WWW), Apr. 2020, pp. 2199–2209.

[26] S. M. Cho, E. Park, and S. Yoo, “MEANITIME: Mixture of attention mechanisms with multi-temporal embeddings for sequential recommendation,” in Proc. 14th ACM Conf. Recommeder Syst., Sep. 2020, pp. 515–520.
YANG XIA was born in 1962. He received the B.S. degree in computer science from Nanjing University, Nanjing, China, in 1984, and the Ph.D. degree in communication and information system from the China University of Mining and Technology (CUMT), Xuzhou, China, in 2009. He is currently a Professor with the School of Computer Science and Technology, CUMT. He has published more than 80 papers in international conferences, international journals, and Chinese journals. His research interests include recommender systems, electronic commerce, and service computing.

YUJING LIU was born in 1996. He received the B.E. degree from the School of Mechanical and Engineering, Beijing Institute of Technology (BIT), Beijing, China, in 2018. He is currently pursuing the M.S. degree with the School of Computer Science and Technology, China University of Mining and Technology (CUMT). His research interests include machine learning and recommender systems.

SONGHE HAN was born in 1995. He received the B.E. degree from the School of Computer Science and Technology, Shanxi Agricultural University (SXAU), Taiyuan, China, in 2019. He is currently pursuing the M.S. degree with the School of Computer Science and Technology, China University of Mining and Technology (CUMT). His research interests include recommender systems and machine learning.

ZIJIE OUYANG was born in 1999. He received the B.E. degree from the School of Computer Science and Technology, China University of Mining and Technology (CUMT), Xuzhou, China, in 2020, where he is currently pursuing the M.S. degree. His research interests include recommender systems and data mining.