Journal of Network Intelligence
Taiwan Ubiquitous Information

Volume **11**, Number **1**, February **2026**

Recommendation of Huizhou Ancient Village Cultural Digital Resources Based on Location-based Social Networks and Attention Mechanisms

Ling Shi[1], Wei-Wei Li[1,*]

[1]The school of arts,
Anhui Xinhua University, Hefei 230088, P. R. China
345124594@qq.com, 564543979@qq.com

Li-Xin Chen[2]

[2]University of Central Lancashire, Preston PR1 2HE, UK
LChen30@uclan.ac.uk

*Corresponding author: Wei-Wei Li

Received September 11, 2024, revised January 24, 2025, accepted April 22, 2025.

Abstract. *In cultural digital resource recommendation, the accurate matching between resources and user interests plays a key role. Traditional recommendation methods ignore the influence of contextual information on the user's choice of resources, resulting in poor recommendation accuracy. For this reason, this paper applies the Location-based Social Network (LBSN) and attention mechanism to Huizhou ancient village cultural digital resource recommendation. Firstly, non-textual information such as user, resource, social, etc. is obtained from the resource access records of LBSN and processed into vectors using Embedding, and GRU is introduced to process the text vectors mapped by BERT model into semantic content vectors, so as to obtain the embedded representation of contextual information of LBSN. Then LSTM is used to gain the short-term predilection of users. Meanwhile, the temporal, spatial and location sequences of user's visits are integrated into the multi-attention scheme to gain users' long-term predilections for resources. Finally, the Common Attention Network (CAN) is adopted to fuse the long and short-term predilections and contextual information, obtain the probability value of each resource by Softmax, and recommend the top N resources to the user. The experimental outcome indicates that the designed approach has a recommendation accuracy of 0.92, which is improved by 0.05-0.19 compared with the other three methods, and can recommend resources to users more accurately.*
Keywords: Location-based social network; Attention mechanism; LSTM; GRU; Cultural digital resource recommendation

---

1. **Introduction.** Huizhou ancient villages are the most vivid display places of Huizhou regional culture. Due to the relative isolation of the natural environment, the historical and cultural landscapes of Huizhou ancient villages have been preserved in the long historical evolution. However, with the rapid socio-economic development, the ancient villages are facing serious challenges because they cannot adapt to the needs of the modern society, and the consequences are often the destruction of the traditional historical style of the ancient villages and the loss of their vitality [1, 2]. Since then more and more ancient villages in Huizhou have begun to enter people's eyes, and while the public awareness of the protection of historical and cultural heritage has increased, people have also become

aware of the value of tourism resources of ancient villages in Huizhou. Under the background of digitalisation, the form of tourism changes, and people can select the cultural digital resources of Huizhou ancient villages through online [3]. However, the vast amount of cultural resources can not accurately match the growing cultural needs of users, how to find and recommend the resources that meet the user's interests in the vast amount of information has been an exigent issue to be addressed nowadays [4, 5]. Therefore, research on the recommendation method of cultural digital resources has far-reaching significance and value for the development of Huizhou ancient villages.

1.1. **Related work.** Early recommendation of cultural digital resources used collaborative filtering (CF) approach, Javari et al. [6] offered a recommendation approach relied on CF that fuses user information with geolocation information, but it needs to manually look for the user's check-in record. Ren and Wang [7] establish geographic, social, and categorical correlations to predict users' relevance scores for unvisited digital resources based on SVM and recommend appropriate resources to users. With the increasing maturity of GPS technology, Location-based social networks (LBSN) have also developed significantly [8]. Ma et al. [9] fused user preference features with geographic information to recommend cultural digital resources to users. Sánchez and Bellogín [10] used the LDA latent probability generating model to integrate social relationships, geographic factors, and other factors for resource recommendation. These CF-based recommendation algorithms incorporate geographic and temporal factors, but they are not expressive enough and suffer from the cold-start problem.

Recently, deep neural networks have brought great performance improvement for digital resource recommendation due to their powerful fitting ability [11]. Khan and Niu [12] suggested a Convolutional Neural Network (CNN) content-enhanced digital resource recommendation model, but the recommendation accuracy is not high. Liu et al. [13] designed a CNN-based sequence characterization approach by treating the check-in sequence as a two-dimensional image and using CNNs for modelling. CNN-based recommendation methods do not take into account the key spatio-temporal contextual factors affecting the user's trips, but Recurrent Neural Networks (RNNs) can solve these problems well and are widely used in modelling sequential data. Ding and Chen [14] proposed to use RNNs to predict the category of the digital resources that a user will visit next, which effectively takes into account the importance of the category attribute for recommendation. Unfortunately, due to the disadvantage of gradient vanishing, RNN can only model short sequences.

LSTM, GRU as a variant of RNN can better model long and short sequences. Lai and Zeng [15] used LSTM to model sequential check-in records to improve recommendation performance. Wang et al. [16] suggested the ST-CLSTM approach, which uses a combination of time gates and distance gates to predict resources that may be of interest to users in the future. Safavi et al. [17] used GRUs to learn the feature representations of check-in sequences within each time window, and combined the resource category information with the feature representations to achieve 83.7% recommendation accuracy.

With deeper research, scholars have found that incorporating the Attention Mechanism (AM) captures the degree of influence of different components to obtain more accurate user preferences. Yang et al. [18] utilized cascading attention to adequately capture users' interests in different dimensions. Cao et al [19] proposed an RNN-AM model for digital resource recommendation, which uses the AM to reflect the importance of the position in the sequence for user signing in. Zhang and Yuan [20] introduced an attention mechanism based on LSTM and used a bilinear matching scheme to calculate the recommendation score for all candidate resources, with a recommendation accuracy of 87%.

1.2. **Contribution.** Although the above research can help users quickly find interested resources from a large number of candidate resources, it ignores the problem of inconsistency of users' short-term and long-term predilections, which leads to unsatisfactory recommendation accuracy. Based on this, this paper designs a digital resource recommendation method for Huizhou ancient village culture based on LBSN and AM. The chief work of the method is summarized as bellow.

(1) Obtain non-textual information from LBSN records and process it into vectors using Embedding, introduce the BERT model to map the resource textual information into a sequence of textual vectors, and process the textual vectors into semantic vectors by using GRU, so as to obtain the embedded representation of contextual information.
(2) The user's most recent sequence of accessed resources is fed into the LSTM to catch short-term resource predilections, and the temporal, spatial, and location sequences of the accesses are then incorporated into the Multiple Attention Mechanism (MAM) to gain the user's long-term preferences.
(3) Adopt a Common Attention Network (CAN), study the user's long-term and short-term predilections and integrate them with the contextual information of the LBSN, output the probability value of each resource through the Softmax function, and recommend the top N resources to the users.
(4) The experimental results on the public check-in dataset in Huizhou District, Huangshan, Anhui Province, show that when the recommended sequence is 20, the Precision, Recall, and NDCG of the suggested model are 0.9409, 0.9251, and 0.5865, individually, which can significantly enhance the accuracy rate of the recommendation model.

2. Theoretical analysis.

2.1. **Location-based social network.** Resource recommendation in LBSN is not only about social relationships between users, but also about location connections, which can be used to recommend resources of interest [21, 22]. As shown in Figure 1, the relationship graph between social, check-in and location on LBSN is illustrated.

Social Relationships: Interactions and connections between users result in different types of social relationships, reflecting similarities in preferences between users.

Check-in relationship: The location that a user checks in can establish a correspondence between the user and the location, which can reflect the user's interest in different locations. The more often a user checks in at a location, the more interested the user is in that location.

Location Relationships: Each location has its own unique latitude/longitude geographic identifier, place name, and location type, etc. The closer the location is to the user's checked-in location or the more similar the location type is, the more likely it is to be checked-in by the user.

2.2. **Recurrent neural network.** RNN as a sequence-to-sequence model dominates the task of recommending cultural digital resources [23]. Unlike BP, CNN, etc., it learns the sequential dependencies of the user's historical accesses using cyclic operations in the internal memory. However, the ability to transfer information is greatly limited as the sequence becomes longer, and the problem of gradient explosion occurs. The variant of RNN, LSTM, GRU can solve the above problems well by introducing the gate structure. For most of the tasks, the performance of GRU is close to that of LSTM. The commonly used formulas in LSTM are as follows.

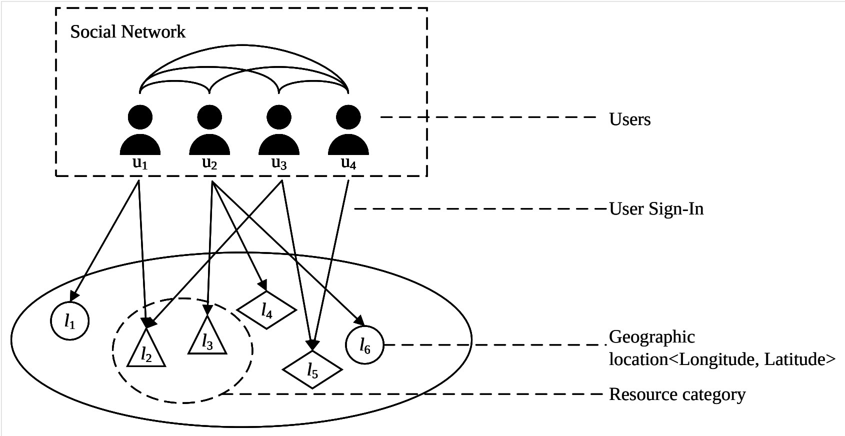

Figure 1. Relationship diagram for LBSN

$$f_t = \delta(V_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

$$i_t = \delta(V_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$o_t = \delta(V_o \cdot [h_{t-1}, x_t] + b_o) \tag{3}$$

where V_f, V_i, V_c, V_o are the weight matrices corresponding to the forgetting gate, the input gate, the intermediate state of neurons and the output gate individually; b_f, b_i, b_c, b_o are the offsets corresponding to the forgetting gate, the input gate, the intermediate state of neurons and the output gate individually; δ represents the activation operation, $h_t = o_t \odot \tanh(c_t)$ represents the obscured state at moment t, and $c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$ represents the neuron state at moment t; x_t represents the input state; $\tilde{c}_t = \tanh(V_c \cdot [h_{t-1}, x_t] + b_c)$ is the intermediate state of the neuron; f_t, i_t, and o_t are the output states of the forget gate, input gate, and output gate, individually; $\odot$ is the dot product operation.

2.3. **Attention mechanism.** The AM learns the sequence information while processing the sequence and focuses more on the crucial information and suppresses or ignores the unimportant information [24], after which better recommendation results can be obtained by continuously adjusting the weight of the information in the sequence. MAM uses different weight matrices on the basis of AM to linearly transform the query to obtain multiple queries, which are computed in parallel by multiple self-attention heads, and then the outputs of all the heads are stitched together to gain the final outcome. The structure of the MAM is indicated in Figure 2, and the specific expressions are as bellow.

$$\begin{cases} \mathrm{MultiHead}(Q, K, V) = \mathrm{Concat}(\mathrm{head}_1, \dots, \mathrm{head}_h) W^O \\ \mathrm{head}_i = \mathrm{Attention}(QW_i^Q, KW_i^K, VW_i^V) \end{cases} \tag{4}$$

where Q, K, V stand for query, robustness and value vectors individually, head stands for attention head, Concat stands for splicing operation, W stands for weights and $Attention(K, V, Q) = \sum_{i=1}^{n} a_i \mathrm{value}_i$, a_i stands for the result of Softmax normalization.

3. **Embedding contextual information embedding for LBSN based on Embedding coding and GRU.** To fully mine the contextual information of the user's accessed resources, the contextual information in the LBSN is categorized into non-textual and textual types [25]. Time series T, geographic location L, social groups G, etc. belong to non-text type, while cultural digital resources are in the form of text, which needs to take into account the information of word order and semantics of the text, while the context

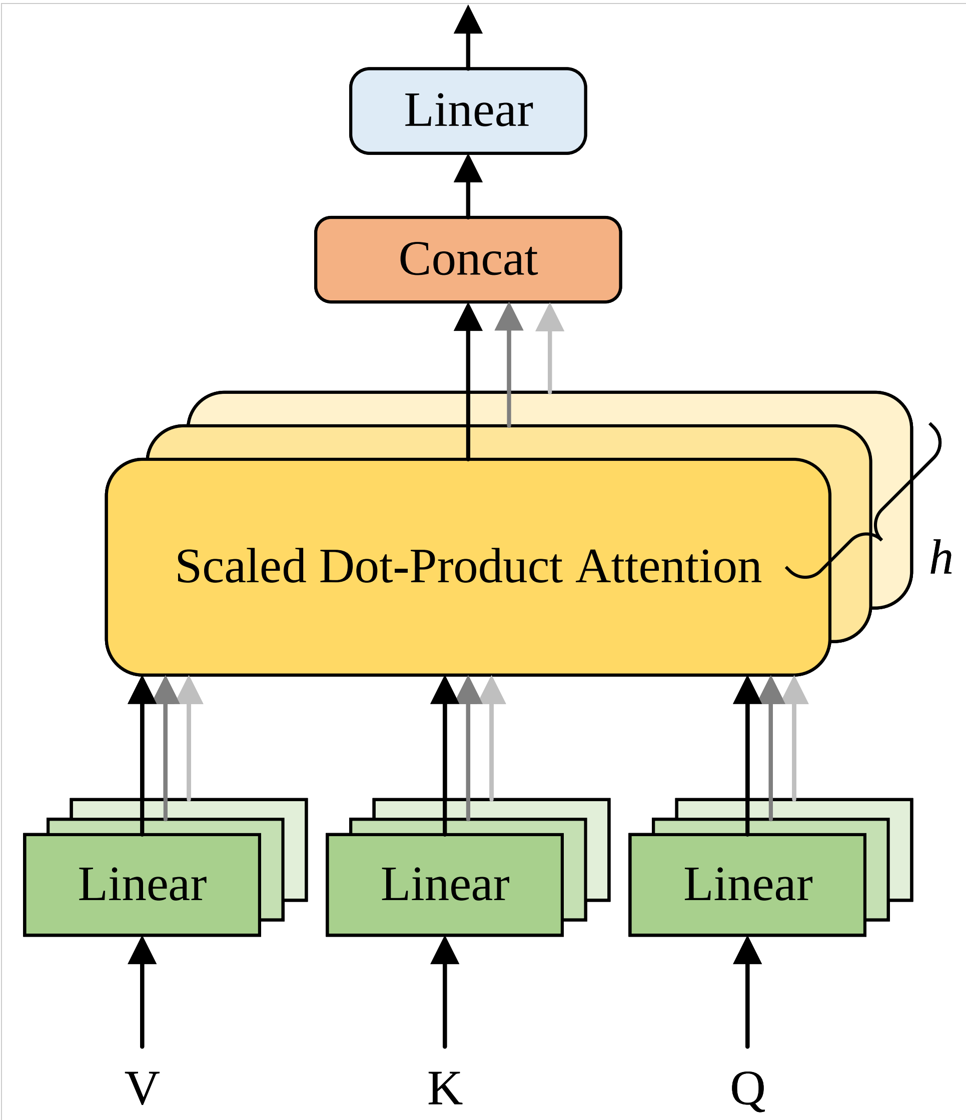

Figure 2. The structure of the MAM

information of the non-text type does not need to be taken into account, so it is essential to consider these two types of data separately.

(1) Non-text type. Since the non-text data in LBSN are discrete types of characters, it is not possible to input the information directly into the neural network for processing. Therefore, discrete symbols or information need to be converted into vectors that can be processed by the neural network. Common vectorization methods include One-Hot coding and Embedding [26], where One-Hot coding results in extremely sparse location matrices, while Embedding uses low-dimensional dense matrices to represent the active locations which is beneficial for neural network processing. The time series T, the geographical location L, and the social group G are processed into feature vectors $T_{emb}^{m\times n}$, $L_{emb}^{m\times n}$, and $G_{emb}^{m\times n}$ using Embedding.

(2) Text types. For a cultural digital resource C containing a large number of text types, the word sequence of the digital resource C is mapped to a sequence of word vectors $C_{BERT} = [w_1, \ldots, w_{t-1}, w_t, \ldots, w_q]$ using the relationship R_{BERT} between words and word vectors trained by BERT [27]. To gain the contextual dependencies in the word sequences and the deeper semantic relationships between word vectors, the reset gate

in GRU is used to adjust the amount of information in the word vector w_{t-1} of the previous word into the candidate state $g_{\tilde{t}}$ of the next word w_t. The reset equation is as follows.

$$r_t = \sigma(a_r g_{t-1} + a_r w_t + b_r) \tag{5}$$

where r_t denotes the size of the reset value at the t-th word in the range $[0, 1]$. w_t denotes the word vector of the t-th word. g_{t-1} is the obscured state of the word vector of the $(t-1)$-th word, and σ denotes the activation operation, as implied bellow.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{6}$$

Multiply the obtained reset gate r_t with the state information g_{t-1} of the word at $t-1$ by elements, splice the result with the current word vector w_t and use the Equation (7) to obtain the candidate state $g_{\tilde{t}}$ of the word vector of the t-th word.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{7}$$

$$g_{\tilde{t}} = \tanh(a_w w_{glove,t} + (r_t \odot g_{t-1})a_g + b_n) \tag{8}$$

Then Equation (8) is used to adjust the amount of information written from the $(t-1)$-th word-vector state g_{t-1} to the t-th word-vector state g_t. The current state g_t consists of candidate states $g_{\tilde{t}}$ and the previous state g_{t-1} dynamically combined in a certain ratio, as shown in Equation (9).

$$z_t = \sigma(a_z g_{t-1} + a_z w_t + b_z) \tag{9}$$

$$g_t = (1 - z_t) \odot g_{t-1} + z_t \odot g_{\tilde{t}} \tag{10}$$

After the processing of the two-layer GRU network to obtain the semantic vector representation of digital resources considering contextual dependencies, word order, and textual semantics $C_{GRU}^{m\times n}$, the final contextual information embedding representation can be obtained as A.

$$A = Concat(T_{emb}^{m\times n}, L_{emb}^{m\times n}, G_{emb}^{m\times n}, C_{GRU}^{m\times n}) \tag{11}$$

4. Recommended digital resources of Huizhou ancient village culture based on LBSN and attention mechanisms.

4.1. **Modelling users' long and short-term predilections based on LSTM and multi-head attention mechanism.** Existing research ignores the problem of inconsistency between users' long and short-term predilections, resulting in poor recommendation outcome. For this reason, after obtaining the embedded representation of LBSN context information, this paper uses LSTM to capture short-term resource preferences while modelling long-term resource preferences using MAM. Then long and short term preferences and contextual information are fused using a CAN to capture dynamic user preferences. Finally, the top N resources are recommended to the user. The framework of the suggested model is indicated in Figure 3.

Considering the problem of inconsistency between users' long-term and short-term predilections, users' predilections for digital resources of Huizhou ancient villages are divided into long-term and short-term predilections. Long-term predilection is the digital resources that users visit frequently, while short-term preference is the resources that users

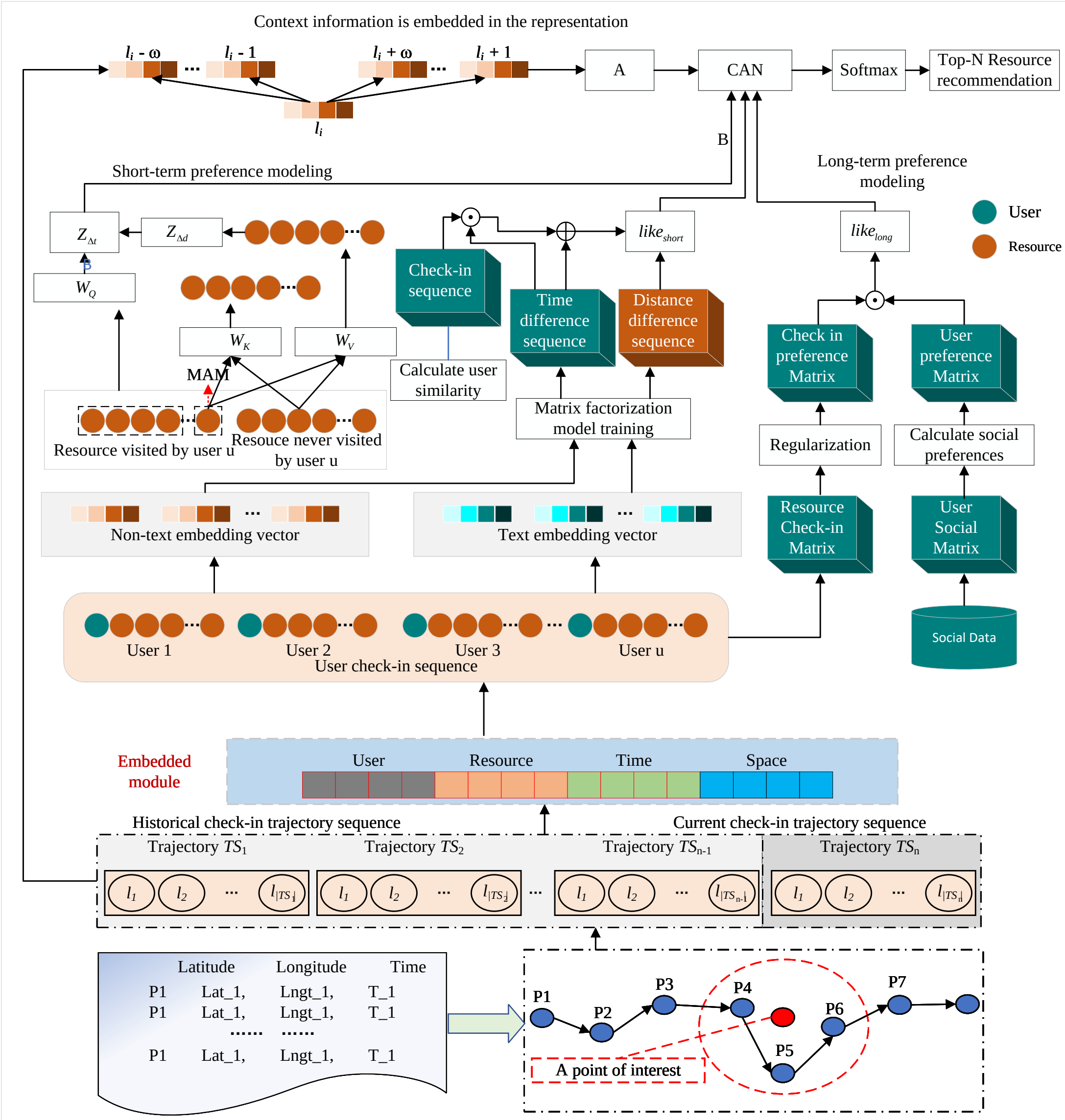

Figure 3. The model structure of the suggested method

browse temporarily. Therefore, LSTM is used to obtain the short-term preference features of users, and MAM is used to learn the location sequences, time difference sequences, and distance difference sequences to obtain the long-term preference features of users.

(1) Short-term preference modelling. Based on the embedding level and the check-in record, the latest M resources $L = \{l_1, l_2, \ldots, l_m\}$ visited by user u can be gained. The resource l_t accessed at the existing moment t and the obscured state h_{t-1} at the moment $t-1$ are added into the neural network layer; the sigmoid operation is used as the activation function to obtain the forgetting vector f_t. All elements of f_t are located in the range of $[0,1]$, and the closer the value of the element is to 0, the more the corresponding information is forgotten, and the closer the value of the element is to 1, the more the information is retained, and the value of f_t is implied as bellow.

$$f_t = \delta(V_f \cdot [h_{t-1}, l_t] + b_f) \tag{12}$$

Input gate i_t selectively remembers the information in l_t and hidden state h_{t-1}, focusing on the important information in l_t and ignoring non-important information, as shown bellow.

$$i_t = \delta(V_i \cdot [h_{t-1}, l_t] + b_i) \tag{13}$$

The unit state update value $\tilde{C}_t$ of the current state is obtained by passing l_t and h_{t-1} through the neural network level with tanh as the activation function, as follows.

$$\tilde{C}_t = \tanh(V_c \cdot [h_{t-1}, l_t] + b_c) \tag{14}$$

The state C_{t-1} at time $t-1$ is multiplied by the oblivion vector f_t at time t and added to the product of the updated value $\tilde{C}_t$ of the state of unit i_t to obtain the state C_t of the unit at time t.

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{15}$$

The digital resources l_t and h_{t-1} are passed by the neural network level with sigmoid as the activation operation to obtain the output vector o_t at time t. Multiplying o_t and C_t to obtain the output h_t. The user's short-term resource preference representation $like_{short}$ can be obtained by executing all the digital resources according to the above process.

(2) Long-term predilection modelling. Long-term predilection is a comparatively stable interest preference, and MAM is used for long-term feature learning due to the large number of user access records and the need to focus on some of the accessed resources while appropriately ignoring non-important resources. To explore the digital resources accessed by users as much as possible, the time difference sequence $\Delta T = [\Delta t_1, \Delta t_2, \ldots, \Delta t_n]$, the coordinate sequence $K = [k_1, k_2, \ldots, k_n]$, and the location sequence $P = [p_1, p_2, \ldots, p_n]$ are extracted from the user's long-term access record $L' = \{l'_1, l'_2, \ldots, l'_n\}$, and the distance difference sequence $\Delta D = [\Delta d_1, \Delta d_2, \ldots, \Delta d_n]$ is computed based on the distance between two latitude and longitude coordinates according to the Equation (16).

$$\Delta d_i = 2arcsin\sqrt{sin^2\frac{a}{2} + cos(k_{ai}) \times cos(k_{ai-1}) \times sin^2\frac{b}{2}} \times R \tag{16}$$

where a denotes the difference in dimension between two coordinate points, b denotes the difference in longitude between the two points, and R denotes the radius of the Earth, 6378.137 kilometres.

The user's long-term visit location sequence P, time difference sequence ΔT, and distance difference sequence ΔD are embedded to obtain P_{emb}, ΔT_{emb}, and ΔD_{emb}. P_{emb}, ΔT_{emb}, and ΔD_{emb} are learnt using the MAM to obtain multiple relevant information in the input sequences.

In P_{emb} set the h-header attention, the sequence length to len and the obscured level size to s. The query matrix $Q^{len\times s}$, the key matrix $K^{len\times s}$, and the weight matrix $V^{len\times s}$ are obtained by computing P_{emb} through h linear transformations of the matrix W_Q, W_K, W_V. Normalizing Q by multiplying it with the K^T matrix yields h weight distribution matrices $M^{len\times len}$. Multiplying M with its corresponding to-be-computed weight matrix V yields h self-attention matrices $Head^{len\times s}$. The above process is shown in Equation (17), Equation (18) and Equation (19).

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{s}}\right)V \tag{17}$$

$$Head_i = Attention(QW_{Qi}, KW_{Ki}, VW_{Vi}) \tag{18}$$

$$Z_p = MultiHead(Q, K, V) = Concat(Head_1, \ldots, head_h) \times W_o \tag{19}$$

According to the above method, feature extraction is carried out on the user's check-in time difference sequence and check-in distance sequence, $Z_{\Delta t}$ and $Z_{\Delta d}$ are obtained. The two are fused with Z_p, and the expression $like_{long}$ of the user's long-term predilection for the activity location is obtained through matrix transformation, and the user's long-term and short-term predilection $B = Concat(like_{short}, like_{long})$ is finally spliced.

4.2. **Incorporating user short and long term preferences and contextual information.** Traditional approaches often use the average obscured state of a recurrent model to denote changes in user predilections and build recommendations by forming a dot product among a user's long and short-term predilection vectors, which fails to identify the importance of an individual user's behavior in deciding which digital resource to visit next. However, whether or not a user will access a modified digital resource also depends on the contextual information in section III, such as temporal, geographic and social factors. This paper uses CANs [28] to fuse long- and short-term predilections and contextual information to catch dynamic user preferences. Specifically, every context is projected to a d-dimensional semantic space, and the following formula is used to compute the weighted attention value H. Subsequently, a latent integrated strategy is used to concatenate the weighted attention values, and a nonlinear connectivity level is used to study the dynamic user's short- and long-term predilections.

$$H_k = \phi(W_k A + b_k) \tag{20}$$

$$\alpha_k = \frac{\exp(v_k^T H_k)}{\sum_j \exp(v_j^T H_j)} \tag{21}$$

$$H = \sum_k \alpha_k H_k \tag{22}$$

where W_k, b_k, v_k and α_k are model parameters and function ϕ is the ReLU activation function.

Given a vector of short-term and long-term preferences B of the target user and a vector of contextual information embedding representations A, the dynamic preferences of the user are computed as follows.

$$a_{joint} = f(W_b B + W_h H + b) \tag{23}$$

$$Score = Softmax(a_{joint}) \tag{24}$$

where a_{joint} is the joint attention score after linking, W_b and W_h are the model parameters, f is the nonlinear function.

4.3. **Model training and recommendation.** After obtaining the dynamic user preferences, this paper adopts the Softmax function to gain the probability value of each resource in the set of cultural digital resources, and recommends the resources with the top N probability values to the users. According to the Softmax function, the conditional probability distribution is defined as follows.

$$p(l_j \mid u_i, t_k) = \frac{\exp(u_i B_j^T)}{\sum_{i=1}^{n} \exp(u_i A_j^T)} \tag{25}$$

Therefore, given the training data set $X = \{u_i, t_k, A_j, B_j\}$, the united probability dispersion can be gained as follows.

$$P_\theta(X) = \prod_{x \in X} p(l_j \mid u_i, t_k) = \prod_{x \in X} \frac{\exp(u_i B_j^T)}{\sum_{i=1}^{n} \exp(u_i A_j^T)} \tag{26}$$

The model parameters $\Theta = \{Q, K, V, W_Q, W_K, W_V, W_o, \lambda_A, \lambda_B\}$, are studied by maximizing the conditional log-likelihood [29] to obtain the following objective function, where ρ denotes the equilibrium factor.

$$L = \sum_{x \in X} p(l_j \mid u_i, t_k) - \rho\Theta \tag{27}$$

It is unrealistic to improve the above target function directly because the computational cost of the full Softmax is proportional to $p(l_j \mid u_i, t_k)$. Therefore, an efficient optimization is carried out by using negative sampling techniques and approximating the initial target L with the target function.

$$L' = \sum_{x \in X} (\log \sigma(u_i A_j^T) + \mu \cdot E_{j \sim P_I}[\log \sigma(-u_i A_j^T)]) - \rho\theta \tag{28}$$

where σ is the activation operation, μ is the amount of negatively sampled resources relied on the noise dispersion, and P_I is the probability dispersion of resource accesses.

5. Experiments and analysis of results.

5.1. Recommended performance comparison. The designed model adopts the public check-in dataset of Huizhou District, Huangshan, Anhui collected by Foursquare as the experimental data. In this paper, the dataset is preprocessed to filter out resources with less than 15 check-ins and less than 10 check-ins, and the summary information of the final dataset is shown in Table 1. In addition, the dataset also contains anonymous user id, resource location id, resource category id, latitude, longitude, UTC time information, and so on. In the experiment, 70% of the dataset is chosen as the training set, 10% as the validation set, and 20% as the test set.

Table 1. Statistical information on Foursquare dataset

Items	**Value**
Number of users	1628
Number of digital resources	7803
Number of sign-ups	30684
Number of social relations	10258

For a fair comparison, an Adam optimizer with a studying rate of 0.001 was used to train all parameters in the model. To hinder overfitting, the regularization factor was set to 5×10^{-5}, the loss rate was set to 0.5, the embedding dimension was set to 350, and the batch size was set to 128. The experimental code was programmed in Python 3.8 development environment, and the deep learning framework adopted was TensorFlow 2.5.0.

This paper adopts Accuracy@N, Precision@N, Recall@N, which are common evaluation metrics in recommendation performance, and NDCG@N, which is a measure of the ranking quality of recommendation models [30], to conduct experiments on the proposed models LBSN-AM, CaSe4SR model [13], RNN-AM model [19], and LSTM-SM model [20] for the experimental Comparison, where N stands for the length of the recommendation list. Table 2 demonstrates the Precision, Recall and NDCG comparisons for various models when N is taken as 5, 10 and 20.

Table 2. Performance comparison of four resource recommendation models

N	**Metrics**	**Model**			
		CaSe4SR	**RNN-AM**	**LSTM-SM**	**LBSN-AM**
5	Precision@5	0.7118	0.7542	0.7833	0.7915
	Recall@5	0.6926	0.7694	0.8059	0.8214
	NDCG@5	0.2029	0.1836	0.2385	0.2617
10	Precision@10	0.7381	0.8452	0.9037	0.9204
	Recall@10	0.7819	0.8157	0.8763	0.9056
	NDCG@10	0.3048	0.3629	0.4066	0.4281
20	Precision@20	0.7435	0.8628	0.9047	0.9409
	Recall@20	0.7629	0.8355	0.9163	0.9251
	NDCG@20	0.4329	0.5016	0.5591	0.5865

When N is 5, the Precision, Recall and NDCG of LBSN-AM are respectively increased by 7.97%, 12.88% and 5.88% compared with CaSe4SR, and by 3.73%, 5.2% and 7.81% compared with RNN-AM. Compared with LSTM-SM, it increased by 0.82%, 1.55% and 2.32% respectively. The overall performance of CaSe4SR is poor due to its inability to learn deep information and its poor ability to learn user preference information. RNN-AM, on the other hand, introduces AM on the basis of traditional RNN to enhance the important digital resources and significantly enhance the learning capability of the model, and thus RNN-AM achieves better performance than CaSe4SR. LSTM-SM does not fully consider social factors and users' long and short-term predilections, resulting in a recommendation performance that is not as good as that of LBSN-AM. Overall, LBSN-AM has some improvement in all the metrics, which verifies the superiority of its recommendation performance.

A comparison of the Accuracy of LBSN-AM with the other three models for various values of N is implied in Figure 4. The accuracy of RNN-AM, LSTM-SM, and LBSN-AM tends to level off gradually as N increases, and the growth rate of the accuracy of CaSe4SR is significantly lower than that of RNN-AM, LSTM-SM, and LBSN-AM. The Accuracy of CaSe4SR and RNN-AM is always lower than that of LBSN-AM with LSTM-SM, while the Accuracy of LBSN-AM with LSTM-SM is almost equal when N is taken from 1 to 5. When N is taken from 10 to 50, the Accuracy of LBSN-AM is significantly higher than that of LSTM-SM, and reaches a maximum of 0.95 when $N = 17$. As N increases, the Accuracy of LBSN-AM gradually stabilizes at 0.92, and the Accuracy of CaSe4SR, RNN-AM and LSTM-SM stabilises at 0.73, 0.85, and 0.87, respectively, and thus the best recommended accuracy is achieved for LBSN-AM.

5.2. **Analysis of ablation experiments.** To further examine the impact of each module in LBSN-AM on recommendation performance, an ablation experiment was conducted, wherein "/EM" represents the removal of non-text data coding module, "/GRU" represents the removal of text semantic embedded representation module, "/SHORT" represents the removal of short-term predilection modeling module, and "/LONG" represents

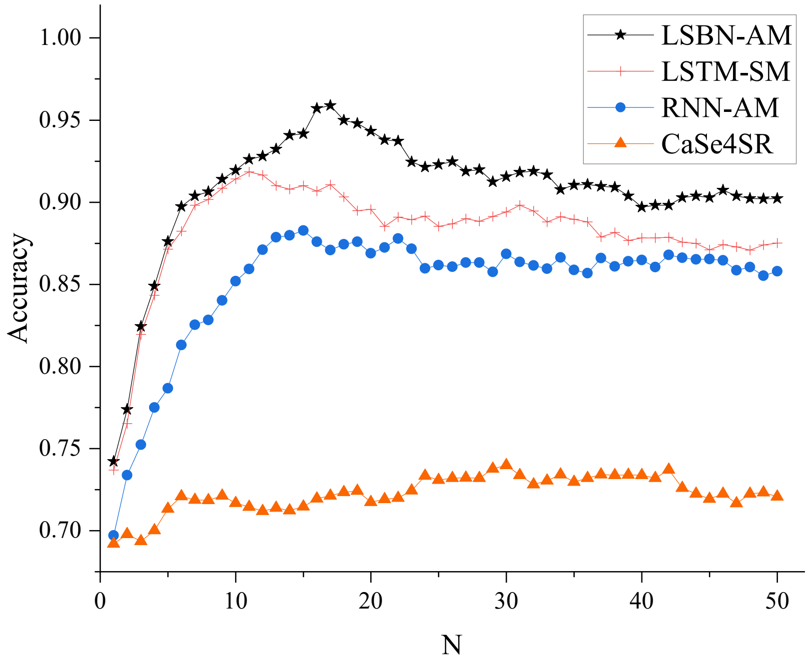

Figure 4. Comparison of accuracy of different recommendation models

the removal of long-term predilection modeling module. "/AM" indicates the use of simple additive fusion instead of attention fusion. When N is 5, 10, and 20, the ablation results of different modules are shown in Figure 5.

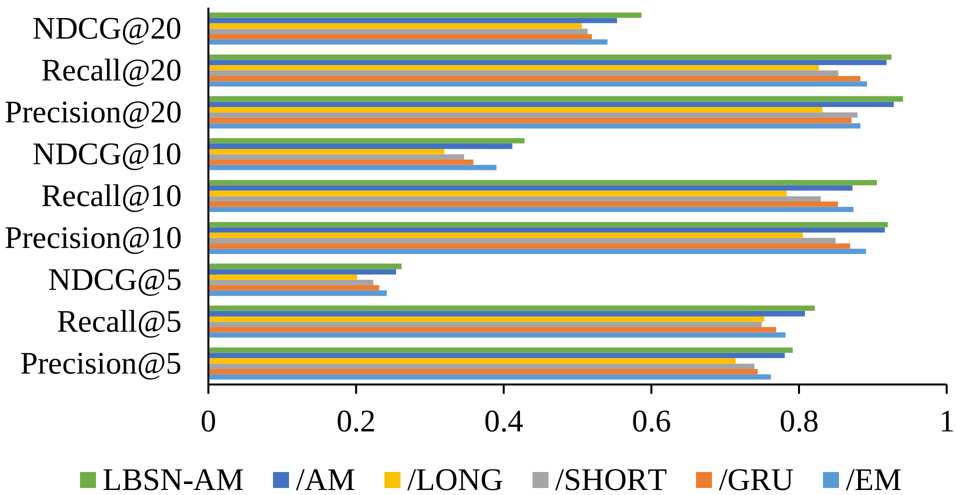

Figure 5. Ablation results of each module in LBSN-AM model

/LONG has the lowest metrics, suggesting that the introduction of a multi-attention mechanism to model users' long-term predilections is critical to the impact of the recommendation model, and /SHORT has the next lowest metrics, suggesting that modelling users' short-term preferences using LSTM is equally important. The metrics of /AM show the smallest decrease compared to the full LBSN-AM, indicating that simple splicing and fusion also affects the recommendation effect. The values of precision@20 of /EM and /GRU are reduced by 5.75% and 7% respectively compared with LBSN-AM, which indicates that Embedding of context information on GRU and GRU can improve the recommendation performance. Therefore, every module in LBSN-AM is indispensable, and LBSN-AM that combines all components is recommended to perform best.

6. **Conclusion.** Existing cultural digital resource recommendation methods ignore the problem of inconsistency between users' long and short-term preferred, resulting in poor matching of recommendations. Intending to this issue, this paper designs a cultural digital resource recommendation method based on location-based social networks and attention mechanism with Huizhou ancient villages as the application object. Firstly, the non-textual information of user accessed resources in LBSN is processed into vectors by using Embedding, BERT is introduced to map the textual information into a sequence of content vectors, and it is processed into textual semantic vectors by using GRU, so as to obtain the embedded representation of contextual information in LBSN. Short-term resource preferences are then mined using LSTM, while long-term resource preferences are modelled using MAM, and a CAN is adopted to fuse long and short-term predilections and contextual information to catch dynamic user preferences. Finally, the probability value of each resource is obtained by Softmax function, and the top N resources are recommended to the user. Experiments indicate that the designed method has high recommendation accuracy and NDCG, and can significantly improve the recommendation matching degree of the recommendation method.

Acknowledgment. This work is supported by the Anhui Provincial Higher Education Scientific Research Project (Philosophy and Social Sciences) (No. 2023AH051804), 2023 Anhui Province Social Science Innovation Development Research project (No. 2023KY504) and Research Center for Quality-oriented Education of College Students 2024 University-level Scientific Research Project (No. IFQE202404).

REFERENCES

[1] Y. C. Zheng, and C. M. Yan, "An exploration of Huizhou architecture connotation in the background of traditional culture changes," *Advanced Materials Research*, vol. 368, pp. 3556–3560, 2012.
[2] T.-Y. Wu, H. Li, S. Kumari, and C.-M. Chen, "A Spectral Convolutional Neural Network Model Based on Adaptive Fick's Law for Hyperspectral Image Classification," *Computers, Materials & Continua*, vol. 79, no. 1, pp. 19–46, 2024.
[3] F. Zhang, T.-Y. Wu, J.-S. Pan, G. Ding, and Z. Li, "Human motion recognition based on SVM in VR art media interaction environment," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, 2019.
[4] Y. Lai, and X. Yang, "Comparative Analysis on the Traditional Villages of Folk Groups of Han Nationality in Huizhou," *Current Urban Studies*, vol. 9, no. 3, pp. 343–354, 2021.
[5] X. Tang, "Under the influence of Huizhou Merchant Culture, Research on Spatial Morphology of Huizhou Ancient Village Landscape Based on Spatial Intelligence Analysis," *Journal of Research Administration*, vol. 5, no. 2, pp. 114–128, 2023.
[6] A. Javari, J. Gharibshah, and M. Jalili, "Recommender systems based on collaborative filtering and resource allocation," *Social Network Analysis and Mining*, vol. 4, pp. 1–11, 2014.
[7] L. Ren, and W. Wang, "An SVM-based collaborative filtering approach for Top-N web services recommendation," *Future Generation Computer Systems*, vol. 78, pp. 531–543, 2018.
[8] J. Bao, Y. Zheng, D. Wilkie, and M. Mokbel, "Recommendations in location-based social networks: a survey," *GeoInformatica*, vol. 19, pp. 525–565, 2015.
[9] Y. Ma, J. Mao, Z. Ba, and G. Li, "Location recommendation by combining geographical, categorical, and social preferences with location popularity," *Information Processing & Management*, vol. 57, no. 4, 102251, 2020.
[10] P. Sánchez, and A. Bellogín, "Point-of-interest recommender systems based on location-based social networks: a survey from an experimental perspective," *ACM Computing Surveys (CSUR)*, vol. 54, no. 11s, pp. 1–37, 2022.

[11] G. C. Marinó, A. Petrini, D. Malchiodi, and M. Frasca, "Deep neural networks compression: A comparative survey and choice recommendations," *Neurocomputing*, vol. 520, pp. 152–170, 2023.

[12] Z. Y. Khan, and Z. Niu, "CNN with depthwise separable convolutions and combined kernels for rating prediction," *Expert Systems with Applications*, vol. 170, pp. 114528, 2021.

[13] L. Liu, L. Wang, and T. Lian, "CaSe4SR: Using category sequence graph to augment session-based recommendation," *Knowledge-Based Systems*, vol. 212, 106558, 2021.

[14] R. Ding, and Z. Chen, "RecNet: A deep neural network for personalized POI recommendation in location-based social networks," *International Journal of Geographical Information Science*, vol. 32, no. 8, pp. 1631–1648, 2018.

[15] Y. Lai, and X. Zeng, "A POI recommendation model for intelligent systems using AT-LSTM in location-based social network big data," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 19, no. 1, pp. 1–15, 2023.

[16] K. Wang, X. Wang, and X. Lu, "POI recommendation method using LSTM-attention in LBSN considering privacy protection," *Complex & Intelligent Systems*, vol. 9, no. 3, pp. 2801–2812, 2023.

[17] S. Safavi, M. Jalali, and M. Houshmand, "Toward point-of-interest recommendation systems: A critical review on deep-learning Approaches," *Electronics*, vol. 11, no. 13, 1998, 2022.

[18] Y. Yang, Y. Duan, X. Wang, Z. Huang, N. Xie, and H. T. Shen, "Hierarchical multi-clue modelling for POI popularity prediction with heterogeneous tourist information," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 4, pp. 757–768, 2018.

[19] G. Cao, S. Cui, and I. Joe, "Improving the spatial–temporal aware attention network with dynamic trajectory graph learning for next Point-Of-Interest recommendation," *Information Processing & Management*, vol. 60, no. 3, 103335, 2023.

[20] Y. Zhang, and Y. Yuan, "Bi-LSTM-AMSM: bidirectional long short-term memory network and attention mechanism with semantic mining for e-commerce web page recommendation," *International Journal of Web Engineering and Technology*, vol. 18, no. 4, pp. 376–396, 2023.

[21] M. J. Chorley, R. M. Whitaker, and S. M. Allen, "Personality and location-based social networks," *Computers in Human Behavior*, vol. 46, pp. 45–56, 2015.

[22] X. Wei, Y. Qian, C. Sun, J. Sun, and Y. Liu, "A survey of location-based social networks: problems, methods, and future research directions," *GeoInformatica*, vol. 26, no. 1, pp. 159–199, 2022.

[23] M. Imran, H. Yin, T. Chen, Q. V. H. Nguyen, A. Zhou, and K. Zheng, "ReFRS: Resource-efficient federated recommender system for dynamic and diversified user preferences," *ACM Transactions on Information Systems*, vol. 41, no. 3, pp. 1–30, 2023.

[24] Y. Zhang, and Q. Ma, "Dual attention model for citation recommendation with analyses on explainability of attention mechanisms and qualitative experiments," *Computational Linguistics*, vol. 48, no. 2, pp. 403–470, 2022.

[25] E. Khazaei, and A. Alimohammadi, "Context-aware group-oriented location recommendation in location-based social networks," *ISPRS International Journal of Geo-Information*, vol. 8, no. 9, 406, 2019.

[26] P. Rodríguez, M. A. Bautista, J. Gonzalez, and S. Escalera, "Beyond one-hot encoding: Lower dimensional target embedding," *Image and Vision Computing*, vol. 75, pp. 21–31, 2018.

[27] C. Mao, J. Xu, L. Rasmussen, Y. Li, P. Adekkanattu, J. Pacheco, B. Bonakdarpour, R. Vassar, L. Shen, and G. Jiang, "AD-BERT: Using pre-trained language model to predict the progression from mild cognitive impairment to Alzheimer's disease," *Journal of Biomedical Informatics*, vol. 144, 104442, 2023.

[28] Z. Miao, X. Zhao, J. Wang, Y. Li, and H. Li, "Complemental attention multi-feature fusion network for fine-grained classification," *IEEE Signal Processing Letters*, vol. 28, pp. 1983–1987, 2021.

[29] X.-B. Jin, C.-L. Liu, and X. Hou, "Regularized margin-based conditional log-likelihood loss for prototype learning," *Pattern Recognition*, vol. 43, no. 7, pp. 2428–2438, 2010.

[30] Y. Zheng, and D. Wang, "Multi-criteria ranking: Next generation of multi-criteria recommendation framework," *IEEE Access*, vol. 10, pp. 90715–90725, 2022.