A Multi-attribute Mining Based Personalized Restaurant Recommendation Method

Kuang Haili, Changliang, Sun Yanpeng, and Bin Chenzhong

***Abstract*—Personalized restaurant recommendation is a hotpot in the field of smart tourism. Most existing methods focus on user preference mining based on user portraits but ignore the importance of mining feature based on integrating multi-source information. Hence, a personalized restaurant recommendation method based on multi-attribute mining (MAM_ResR) is proposed. Having constructed restaurant knowledge graph, combining geospatial semantics obtained from dining sequence trajectory and inherent attribute semantics mined by network embedding of restaurant-attribute graph extracted from restaurant knowledge graph, features of restaurants fusing multi-source information are obtained. Then, using explicit feedback and dining trajectory, the user preference fusing multi-attribute feature is obtained. Finally, the recommendation list is calculated by similarity between restaurant characteristics and user preferences. Experiments based on real data show that it is very effective to integrate multiple types of information into the recommendation method to improve recommendation performance.**

***Index Terms*—Multi-attribute mining, sequence representation learning, network representation learning, explicit feedback, personalized restaurant recommendation.**

I. Introduction

With the rapid development of tourism, tourists can make dining plan in advance based on information on social media or vertical domain websites. The vertical domain website not only presents inherent attribute information of restaurants, but also shows related users' feedback information such as ratings and evaluations, which can be reference to others. However, in order to alleviate the drawbacks caused by information overload, it's of great importance to develop a personalized restaurant recommender, which can effectively mine multi-attribute information for restaurant feature learning and user preference learning [1].

At present, the recommendation systems are mainly divided into the following categories: association rules [2], content-based recommendation [3], collaborative filtering [4], and mixed recommendation [5]. The core problem of personalized recommendation is how to effectively integrate user's historical data, item attributes, context information and mine user preferences. Therefore, comprehensive utilization of available information is the future trend of recommendation system [6]. Deep learning has been widely used, combining which with recommendation algorithm has become a research hotpot in recent years. Compared with the traditional recommendation method, these methods have obvious advantages [7], which can effectively extract item features automatically, and can fully mine user preferences and interests to achieve user modeling.

In this paper, considering mining multi-source information with deep learning method for recommendation, we propose a multi-attribute mining based personalized restaurant recommendation model (MAM_ResR). Since the geospatial semantic information of restaurant is hidden in the user's dining sequence, we first use the distributed representation method to model dining sequence to obtain geospatial semantic feature. Then we use the network representation learning method to mine attribute feature of restaurant in restaurant-attribute network extracted from restaurant knowledge graph. Combining both, we can obtain vector representation of restaurant fusing multi-source information. Regarding rating score as attention weight of each restaurant in dining trajectory, we can obtain user preference feature. Finally, personalized restaurant recommendation can be achieved based on the vectors similarity between restaurant and user. Experiments on real dataset collected form Dianping show the effectiveness of our model in restaurant recommendation task.

The main contributions of this paper are as follows:

1) We construct the knowledge graph of Guilin restaurants. According to restaurant and inherent attribute information clawed from Dianping, the knowledge graph of Guilin restaurant domain is constructed. Based on this knowledge graph, restaurant-attribute network can be extracted.

2) We propose a multi-source feature leaning method based on multi-attribute mining method. Considering one-sideness of modeling dining trajectory for geospatial feature semantic, we fuse inherent attribute semantic obtained from restaurant-attribute network embedding to achieve multi-source feature learning.

3) We propose a personalized restaurant recommendation method based on fusing multi-attribute mined from historical dining trajectory, restaurant-attribute network and rating feedback information. In this paper, we use datasets crawled from websites in vertical domain, such as Dianping, to verify that the proposed method has good practicability.

The rest of this paper is organized as follows: after describing the existing works related to the current research in Section 2, we give the problem description and system framework and introduce the MAM_ResR model based on multi-attribute mining in detail in Section 3. Section 4 represents the data processing work and the experimental results of the model on the real dataset. Section 5 summarizes the content of this article and puts forward suggestions for

Manuscript received January 20, 2019; revised June 12, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant U1711263 and Grant U1811264, in part by the Guangxi Innovation Drive Major Special Projects Under Grant AA17202024, and in part by Innovation Project of GUET Graduate Education Under Grant 2018YJCX52.

The authors are with the Guilin University of Electronic Technology, Guilin 541004, China (e-mail: 925211089@qq.com, changl@guet.edu.cn).

doi: 10.18178/ijke.2019.5.1.109

future research directions.

II. Related Work

A. Sequence Representation Learning

There is a strong correlation among local data of a sequence. By mining the association among local data and using sequence representation learning method, not only the individual's feature semantics in the sequence can be obtained, but also the whole sequence's feature semantics can be obtained. Sequences exist in many forms, such as text, dining trajectory, and product click records.

Word2vec [8] is an efficient tool to represent words as real-valued vectors proposed by Google in 2013. This tool uses the idea of deep learning to simplify the processing of text content into vector operations in k-dimensional vector space through training, and expresses the semantic similarity of text by similarity in vector space. Therefore, many researchers attempt to solve the problem of trajectory sequence modeling based on the idea of word2vec [9].

In summary, sequence representation learning can capture the implicit semantics between items in trajectory sequence well [10]. It has strong practicability in application [11]-[13]. However, the semantics of the sequence is limited, and other auxiliary information needs to be fused to improve the recommendation performance.

B. Network Feature Embedding

Practical recommendation algorithms need strong scalability, which can easily integrate other auxiliary information. Information with network structure is special. In recent years, network feature learning [14] has gradually become a hot research direction in machine learning. The basic idea of network representation learning is to maintain the similarity of nodes in embedded space and enhance the performance of network mining task [15]. Perozzi [16] proposes a feature learning algorithm using deep learning for graph structure learning. Under the guidance of Word2vec, it introduces random walk to generate sequence for training. Grover [17] introduced in 2016 is inspired by Word2vec and Deepwall. It is an algorithm that maps every node in the network to a vector in low-dimensional space by unsupervised learning method. This algorithm returns a feature representation [18] that maximizes the possibility of preserving the network neighborhood of nodes in low-dimensional feature space. Because of a large number of information with network structures, combining network feature learning with recommendation system [19]-[21] can effectively enhance the learning ability of recommendation system and improve the accuracy and user satisfaction of recommendation system.

III. The Multi-attribute Mining Based Personalized Restaurant Recommendation Model (MAM_ResR)

For the convenience of reading, we first introduce the structure and definitions used in some models.

***Definition 1*: Dining Check-in.** In social networking service, each check-in matches a user, a restaurant and a time stamp. The form of check-in data d_s is: $<u, t, r>$. Here is a dining check-in example *<Li, 2018/06/12, KFC >*, denoting that the user whose id is *Li* has dined at the restaurant named *KFC* on the day of *June 12th, 2018*.

***Definition 2*:** Dining sequence. We use the item "dining sequence" to denote user's dining situation ordered by time stamp, the form of which is: $<u, t_1, r_1>, \ldots, <u, t_i, r_i>, \ldots, <u, t_N, r_N>$, where N denotes the length of series and $t_i < t_{i+1}$ *for* $i \leq N-1$.

***Definition 3*:** Restaurant-attribute network. The restaurant-attribute network can be represented as an undirected graph of $G=(V, E)$, where V represents nodes in the network, E represents a undirected edge connecting nodes. A restaurant or a specific attribute is a node in the network. Edge joints the intrinsic attribute information and a restaurant.

***Definition 4*:** Dining feedback on rating. In relative service on vertical websites, like Dianping and Meituan, a user can explicitly express a feedback to the restaurant the user has consumed by a rating score from 1 to 5, which indicates attitude from negative to positive.

***Problem formulation*:** In this study, we consider the restaurant recommendation problem as follows: given a collection of users' dining check-ins, users' ratings to restaurants and restaurants' attributes (e.g., star-level, mean-price, taste-level, etc.), we aim to recommend each user with a list of candidate restaurants.

Framework of the proposed model is shown in Fig. 1. It mainly consists of three modules: (1) vector representation learning of restaurant geospatial feature semantics (2) vector representation learning of restaurant inherent attribute feature semantics and (3) user preference mining with explicit feedback as attention weight. The three modules are discussed separately in the following sections.

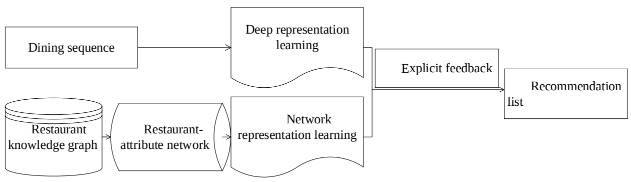

Fig. 1. MAM_ResR model framework.

A. Restaurant's Geospatial Semantic Feature Vector Learned from Dining Sequences

The user's dining trajectory has a high similarity with the text data. In the text document, the appearance of each word is affected by its context. In the user's dining trajectory, each dining behavior reflects the influence of past and future dining behaviors. Inspired by the word2vec idea, we model the dining trajectory sequence to obtain restaurant's geospatial semantic feature vector.

Given the dining trajectory, $t = \langle (u, t_1, r_1), \ldots, (u, t_n, r_n) \rangle$, let probability $p(r_1, r_2, \ldots, r_n)$ be the occurrence rate of this trajectory, we formulate the probability as follows:

$$p(r_1, r_2, \ldots, r_n) = p(r_1) \times p(r_2 | r_1) \times \ldots \times p(r_n | r_1, r_2, \ldots, r_{n-1}) \quad (1)$$

Assuming that the appearance of each restaurant is only related to surrounding 2C restaurants, then:

$$p(r_1, r_2, \ldots, r_n) = \prod_{i=1}^{n} p(r_i | context_{r_i}) \quad (2)$$

where $context_{r_i}$ represents the former C and later C restaurants adjacent to the restaurant. A window with size of C slide from beginning to end, and the words on both sides of the window can be guessed from the middle words of the window. The change of C can control the number of words in the context window. In the training process, selecting a small number of negative cases can effectively reduce the update amount of weight parameters and greatly improve the calculation efficiency. The parameter controlling the number of negative cases is negative.

Assuming that there are *N* trajectories in the trajectory set and *NT* restaurants in time order are included in the trajectory *T*, the objective function of the model is:

$$L = \prod_{T=1}^{N} \prod_{i=1}^{N_T} p(r_i | context_{r_i}) \quad (3)$$

User dining trajectory sequence is used to train word2vec three-layer neural network model, update parameter [22], obtain restaurant geospatial semantics, and get restaurant low-dimensional dense vector representation. The results are the basis for subsequent operations.

B. Restaurant's Inherent Attribute Semantic Feature Vector Learned from Knowledge Graph

Considering that the dining trajectory sequence can only obtain the geospatial semantics of the restaurant, but ignores the inherent attribute semantics. Extracted from constructed restaurant knowledge graph, inspired by node2vec, the restaurant-attribute network can be mapped into the low-dimensional vector space for attribute semantics mining. A restaurant or a specific attribute is a node in the network. Adding all attributes vectors of a restaurant together, we can obtain restaurant inherent attribution semantics.

The extraction of network features is to learn the characteristics of the network by solving the optimization function. This process can be regarded as the maximum likelihood optimization problem. The node2vec method defines the representation of the node network and proposes a random walk strategy with bias to generate sequence, which can effectively retrieve the scattered neighboring nodes and return the feature representation of the network in a low-dimensional space. These feature expressions maximize the likelihood estimation of the neighboring nodes of the node.

Given a network, G=(V,E), the projection function that maps the nodes in the network to the feature space is $f: V \to R^d$, where d represents the vector dimension, that is to say, f is a matrix of size being $|V| \times d$. For any node u in the network, $u \in V$, the neighbor node set, $N_S(u) \subset V$, is generated by neighbor node sampling method *S*. The loss function is defined as:

$$\max_f \sum_{u \in V} \log P_r(N_s(u) | f(u)) \quad (4)$$

In order to facilitate calculation, two hypotheses are introduced:

1) Conditions are independent, and each neighbor is independent to each other:

$$P_r(N_s(u) | f(u)) = \prod_{n_i \in N_s(u)} P_r(n_i | f(u)) \quad (5)$$

2) The symmetry of feature space, in which the influence of source node and neighbor node on each other is the same:

$$P_r(n_i | f(u)) = \frac{\exp(f(n_i) \cdot f(u))}{\sum_{v \in V} \exp(f(v) \cdot f(u))} \quad (6)$$

By synthesizing the above three formulas, the final results to be optimized are obtained:

$$\max_f \sum_{u \in V} \left[-\log Z_u + \sum_{n_i \in N_S(u)} f(n_i) \cdot f(u) \right] \quad (7)$$

For any node in the network, $u \in V$, with Node2vec method, Num-walks sequences whose length is walk-length are generated. Assuming that currently residing on node *v*, the last step is to transfer from node *T* to node *v*, which represents the shortest path distance between node *T* and node *x*, then the next sampling node *x* is determined by probability distribution:

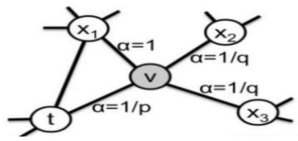

Fig. 2. Random walk strategy.

$$\alpha_{pq}(t, x) = \begin{cases} \frac{1}{p} \text{ if } d_{tx} = 0 \\ 1 \text{ if } d_{tx} = 1 \\ \frac{1}{q} \text{ if } d_{tx} = 2 \end{cases} \quad (8)$$

where *p* is the return probability parameter and *q* is the departure probability parameter. The significance of both is to control the balance between depth priority and breadth priority. In order to obtain the optimal values of hyper parameters *p* and *q*, node2vec uses the semi-supervised form to learn the node representation using the grid to search for the most appropriate parameters.

The restaurant-attribute graphs are trained by node2vec to obtain the inherent attribute semantics of the restaurant, and each attribute vector of each restaurant is obtained. The attribute feature vector at this time represents a vector representation corresponding to each attribute value, and a vector representation of the intrinsic attribute feature semantics of each restaurant is obtained by accumulating various attribute value vectors of the respective restaurants.

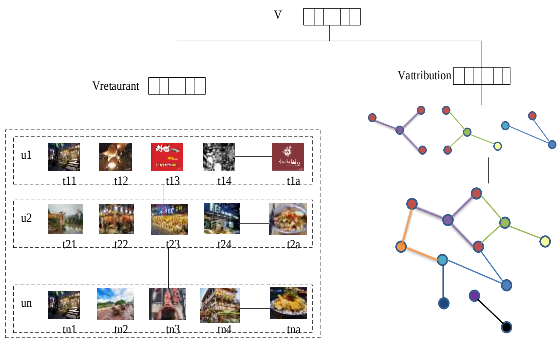

Fig. 3. Vector representation learning of restaurant.

Finally, by adding the geospatial semantic feature vector and the intrinsic attribute semantic feature vector of the restaurant, the restaurant vector representation fusing multi-source information is finally obtained, that is, $V = V_{restaurant} + V_{attribution}$.

C. User Dining Preference Learned by Hard Attention Method

The vector representation of the restaurant is obtained by combining the geospatial semantic features and inherent attribute semantic features. In traditional methods, user vector representation is gained by modeling human behavior and generating user portrait, which ignores the information hidden in explicit feedback. Considering the idea that user dining preference can be discovered from rating score, we propose a hard attention method to obtain the vector representation of user dining preference.

The explicit feedback, rating score, directly indicates the user preference for the restaurant. Inspired by the attention mechanism that when people notice a target or a scene, the distribution of attention within the target or at every spatial location in the scene is different, we apply rating to being the corresponding attention for a restaurant. The higher the score, the more the restaurant contributes to the recommendation result, and there should be more attention paid to this restaurant. Otherwise, lower rating score means the user's less attention to a restaurant. Therefore, with user dining trajectory, each restaurant should be given a corresponding degree of attention according to the rating feedback given by the user for different restaurants.

Given user u with check-ins $<u, t_1, r_1>, \dots, <u, t_i, r_i>, \dots, <u, t_N, r_N>$, dining historical sequences with rating scores can be listed as $\{<r1, s1>, \dots, <ri, si>, \dots, <rN, sN>\}$. Using restaurants vectors obtained from network embedding, with rating regarded as corresponding attention, users vectors can be seen as a multiplication result.

$$u_i = \sum_{j=1}^{n} \frac{V_j * \omega_{ij}}{n} \qquad (9)$$

where u_i means vector representation of the i^{th} user, n denotes number of restaurants, V_j means vector of the j^{th} restaurant, and ω_{ij} denotes the attention of user i's rating score on restaurant j, and ω_{ij}=softmax(r)=$\frac{exp(\theta_i^T x)}{\sum_{k=1}^{K} exp(\theta_k^T x)}$.

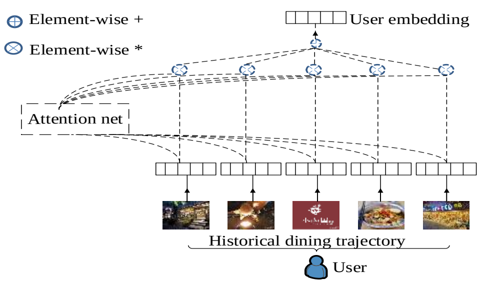

Fig. 4. Vector representation learning of user dining preference.

D. Model Application

We will apply the proposed model to the real restaurant recommendation task. In the real restaurant recommendation task, the user is more likely to select restaurants similar to the previously visited restaurant with high raring score. Therefore, we generate a TOP-K recommendation list according to the recommendation formula $S(u, r)$:

$$S(u, r) \propto (V_{restaurant} + V_{attribute})^T \cdot u \qquad (10)$$

where u represents a attention-based semantic feature vector of user u, $V_{restaurant}$ represents the geospatial semantic feature vector, and $V_{attribute}$ devotes the attribute semantic feature vector.

IV. Experiment

In this section, based on real datasets, we evaluate the proposed model. Firstly, the data processing and the settings in the experiment are introduced. Then, the evaluation criteria are introduced. Finally, the experimental results of the model are evaluated by the evaluation criteria and compared with other methods.

A. Data Preprocessing

In order to investigate and evaluate the preferences of the eating of users, we collect 25267 check-in and rating of users from the public comments. As the largest crowdsourcing website for local businesses in China, the website focuses on the catering industry in China and has a broad mass base. The digital footprint, such as check-in, score and comment, usually represents the preferences of users for restaurants.

Trajectory construction: Sorting out the check-in data of user according to the time order, the eating trajectory of user can be obtained. In order to remove the noise, the check-in data of user in the sequence can be filtered and deleted less than three times.

Knowledge Graph Construction: Restaurant knowledge graph stored in Neo4j graph database can be constructed by restaurant and attribute information.

Restaurant-Attribute Network Construction: Restaurant-attribute network under each attribute is extracted and constructed separately according to the attributes from the knowledge graph constructed, forming a network structure under a single attribute that each restaurant is connected with the corresponding attribute values. A total of seven restaurant-attribute network structures are generated, namely: restaurant-environment attribute network, restaurant-per capita consumption attribute network, restaurant-service attribute network, restaurant-taste attribute network, restaurant-star attribute network, restaurant-category attribute network, restaurant-geographic location attribute network.

The construction process of the restaurant-attribute network is shown in Fig. 5.

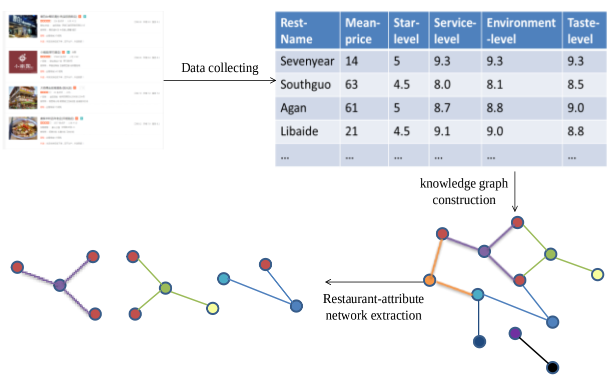

Fig. 5. Restaurant-attribute network extracting process.

B. Experimental Setup

Firstly, the parameters of the contrast algorithm are adjusted, and the effects of different parameters of the algorithm on the experimental results are observed through many experiments. Table I lists the adjustable parameters and meanings.

TABLE I: LISTS AND MEANING OF PAREMETERS

Parameter	Parameter meaning
K	Recommended list length
window_size	Sliding window size
negative	Negative number
num_walks	Node
walk_length	Random walk sequence length
p,q	Balance depth-first and breadth-first parameters

In the experiment, K is set to be 10, and the parameters of the model window=8, negative=4, walk-length=110, num-walks=18, dimension=200 are elaborated through experimental comparison results.

C. Evaluation Criteria

The recall rate index commonly used in recommendation system is used for evaluation. In order to ensure the stability of the index, 20 experiments are needed to adjust the parameters each time, and the average of 20 experiments is taken as the final result.

Randomly select 20% of the data in the database as the testsets. Using our model, a top-K candidate restaurant list is calculated by ordering similarity between restaurant feature semantics and user preference semantics. Definition hit_r@k for a single test user, the recall rate for that user is:

$$hit_r@k = \frac{k_i}{TP_u} \tag{11}$$

where $k_1$denotes the number of restaurants included in the recommended top-*K* attraction that existed exactly in the user u history dining trajectory, and TP_u is the total number of restaurants that user u has visited.

The recall rate Recall@k for all test users indicates the probability of a successful prediction for all test sets:

$$\text{Recall@k} = \frac{\sum hit_r@k}{\#all\_test\_user} \tag{12}$$

D. Comparison Algorithm

In order to verify the effectiveness of fusing multi-source information for enriching the semantics of item features and user preferences mining based on user's display feedback, two algorithms are selected to compare: Doc2vec [23] trains user's dining sequence with word2vec three-layer neural network model, updates parameters, obtains the feature vectors of user's dining trajectory, and obtains the low dimension vector representation of restaurant in vector space. This method only mine the spatial feature semantics, and user preferences are represented by the sequence features of their historical dining trajectory. TSTKG4Rec [24] mine not only the geospatial semantic features of restaurants from dining trajectory but also inherent attribute feature semantics from restaurant-attribute network, and the user preferences are expressed by using the sequence features of users' historical dining trajectory which regard equal attention weight to restaurant in dining trajectory.

E. Analysis of Experimental Results

1) Effect of different K on the recommended effect

The parameter in the experiment: K, which is the length of the recommended list. The performance of the above three algorithms under different K values is tested, wherein the value range of K is [1], [20].

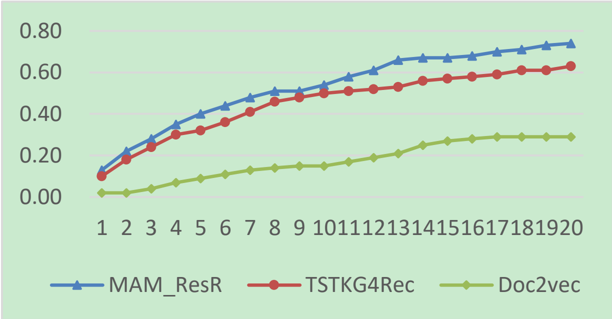

Fig. 6. Recall rate of the three methods.

Fig. 6 shows the overall recommended effects of the three algorithms under different K values. The analysis concludes that the recommended effects of each model increase with the increase of K value. Using this dataset, compared with TSTKG4Rec, the recommended performance of the proposed algorithm is increased by 17% on average, reflecting that user preference information can be reflected by user feedback to a certain extent. In general, the proposed model has a better recommendation than the comparison algorithm, indicating that the proposed MAM_ResR model is more prominent in the recommendation performance than the other two comparison algorithms in the restaurant recommendation. It can be well verified that compared with the Doc2vec method of obtaining the geospatial semantics of restaurants by using the dining trajectory alone and the TSTKG4Rec method of acquiring the inherent attribute semantics of the restaurant by using the knowledge graph alone, combining multi-attribute is effective to improve the recommendation effect. The method proposed in this paper not only explores the geospatial semantics of restaurants in the dining track, but also mines the inherent attribute semantic information of the restaurant hidden in the restaurant-attribute network. It uses the restaurant vector that integrates multi-source information and combines the user's explicit feedback to mine user preference fusing multi-source information. The recommended effect of this paper reflects that it is very effective to integrate the useful information into the recommendation method to improve the recommendation performance.

2) Effect of different p, q on recommendation effect

We have introduced p, q to sample neighbor nodes in the above chapter. Using grid search to adjust them, selecting different p and q combination values can affect the ratio of homogeneity information to isomorphism information in the walk sequence. The effect of different p, q combination values on the experimental results is shown in Table II.

It can be found from Table II that the proposed algorithm is significantly superior to TSTKG4Rec in terms of recommendation performance. When the optimal values of p and q are 2, 3, the experimental results are the best. The significance of p and q is to control the balance of deep traversal and breadth traversal during random walk. The experimental results show that under the data set, when p=2 and q=3, the structural properties and contents in the sequence generation process are similar. The degree reached the most balanced state.

TABLE II: EFFECT OF P, Q ADJUSTED ON THE IMPROVEMENT OF RECOMMENDATION

p	q	ours	node
1	1	0.53	0.53
1	2	0.54	0.48
1	3	0.49	0.43
2	1	0.46	0.37
2	2	0.51	0.48
2	3	0.59	0.5
3	1	0.57	0.44
3	2	0.55	0.47
3	3	0.51	0.45

3) The effect of different context window size values on recommended

The parameter in the experiment: window-size, which is the size of the context window during training.

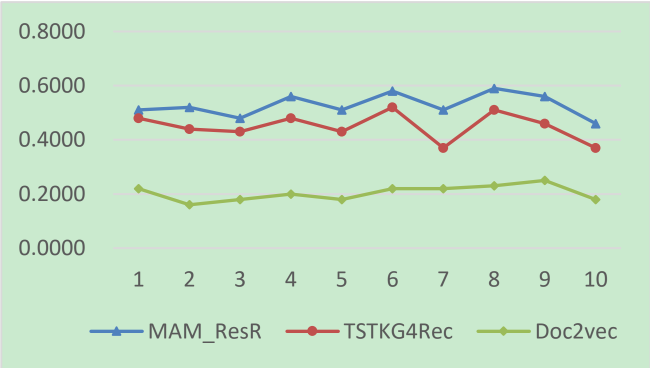

Fig. 7. The effect of Window_size adjusted on the promotion of recommended.

As can be seen from the results shown in Fig. 7, taking different window-size values has different effects on the recommendation results and it's not linear. Performance of the proposed model under different window-size values is significantly better than the other two comparison algorithms. When the window-size value is 8, the proposed algorithm and the comparison algorithm achieve the best results. Compared with TSTKG4Rec, the recommended performance of the proposed algorithm is increased by 18% on average.

4) Effect of different negative sampling values on the recommended effect

The parameter in the experiment: negative, used to set the number of negative cases selected during the training.

Fig. 8. The effect of Negative adjusted on the improvement of recommendation.

As can be seen from the results shown in Fig.8, the adjusting of the negative value has influence on the recommendation result. From the overall effect, the performance of the proposed model under different negative values is significantly better than the other two comparison algorithms. When the negative value is 4, the proposed algorithm and the comparison algorithm achieve the best results.

5) Effect of different sequence lengths on recommended effects

The parameters in the experiment: num-walks, that is, the number of random walk paths through the node.

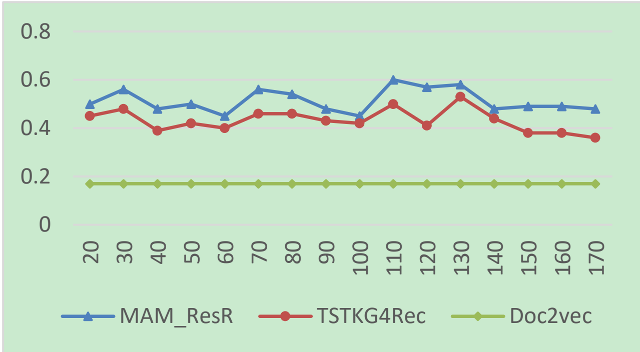

Fig. 9. The impact of num-walks adjusted on the promotion of recommendations.

It can be seen from Fig. 9 that when the number of random walk sequences passing through the node is 18, the restaurant inherent attribute semantics in the restaurant knowledge map can be mapped to the low-dimensional space to the maximum extent. Fewer sequence paths can't fully obtain network feature. If the number of random walk paths of nodes is too large, it is easy to generate noise and cause semantic loss.

6) The effect of different sequence lengths on recommended

The parameter in the experiment: walk-length, which is the length of the random walk sequence.

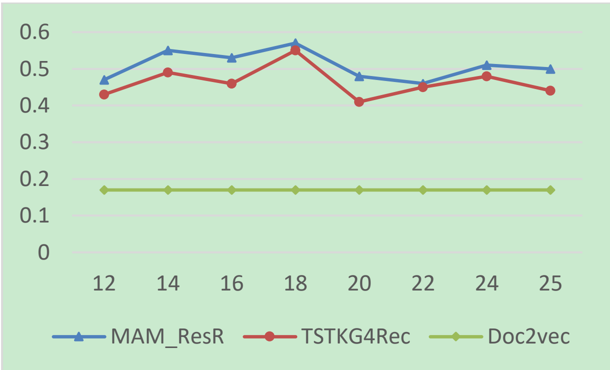

Fig. 10. The effect of adjusted walk-length on the improvement of recommendation effect.

From the results shown in Fig. 10, it can be seen that the recommended performance of the proposed model under different walk-length values is significantly better than that of the other two comparison algorithms. When walk-length value is 110, the proposed algorithm and TSTKG4Rec achieve the best results by setting sliding windows size as 4. Compared with the recommendation effect in doc2vec algorithm with fixed sequence length of 30, the proposed algorithm and the recommendation effect in TSTKG4Rec change with the change of random walk sequence length, but the recommendation performance is always better, which shows that the network feature semantics is easier to mine than the sequence feature semantics.

V. Conclusion

In this paper, a personalized recommendation method based on multi-attribute mining is proposed. This method combines geospatial semantics obtained from the dining trajectory with inherent attribute semantics obtained from restaurant knowledge graph to mine multi-attribute restaurant feature. Combining explicit feedback and dining trajectory, user preference feature are mined. The recommendation list is obtained by calculating the similarity between feature vector representation of the user and the restaurant. In order to evaluate the proposed algorithm, this paper introduces two kinds of experiments. Confirming the appropriate parameter settings, and comparing the proposed method with other two methods, the experiments verified that the recommendation method fusing more auxiliary information can improve the recommendation performance. Explicit feedback is effective in mining user preference. The next work is to further consider how to combine user implicit feedback, such as comment information, to obtain user preferences with richer semantics, so as to improve the recommendation performance of the algorithm.

References

[1] F. Zhang, N. J. Yuan, and Z. Kai, "Exploiting dining preference for restaurant recommendation," in *Proc. International Conf. on World Wide Web*, 2016.

[2] S. W. Changchien and T. C. Lu, "Mining association rules procedure to support on-line recommendation by customers and products fragmentation," *Expert Systems with Applications*, vol. 20, no. 4, pp. 325-335, 2001.

[3] H. Gao, J. Tang, and H. Xia, "Content-aware point of interest recommendation on location-based social networks," in *Proc. Twenty-ninth Aaai Conf. on Artificial Intelligence*, 2015.

[4] Z. D. Zhao and M. Shang, "User-based collaborative-filtering recommendation algorithms on hadoop," in *Proc. 3rd International Conference on Knowledge Discovery & Data Mining*, 2010.

[5] M. Liu, W. Pan, and L. Miao, "Mixed similarity learning for recommendation with implicit feedback," *Knowledge-Based Systems*, vol. 119, no. C, pp. 178-185, 2017.

[6] S. Jiang, X. Qian, and T. Mei, "Personalized travel sequence recommendation on multi-source big social media," *IEEE Transactions on Big Data*, vol. 2, no. 1, p. 1, 2016.

[7] W. Jian, J. He, and C. Kai, "Collaborative filtering and deep learning based recommendation system for cold start items," *Expert Systems with Applications*, vol. 69, pp. 29-39, 2017.

[8] T. Mikolov, I. Sutskever, and C. Kai, "Distributed representations of words and phrases and their compositionality," *Advances in Neural Information Processing Systems*, vol. 26, pp. 3111-3119, 2013.

[9] Y. Kim, Y. Jernite, and D. Sontag, "Character-aware neural language models," *Computer Science*, 2016.

[10] S. Zhang, L. Yao, A. Sun, *et al. Deep Learning Based Recommender System: A Survey and New Perspectives*, 2017.

[11] F. Zhuang, Z. Zhang, and M. Qian, "Representation learning via dual-autoencoder for recommendation," *Neural Netw*, vol. 90, pp. 83-89, 2017.

[12] O. U. Florez, "Deep learning of semantic word representations to implement a content-based recommender for the RecSys challenge'14," *Communications in Computer & Information Science*, vol. 475, pp. 199-204, 2014.

[13] S. Gupta and V. Varma, "Scientific article recommendation by using distributed representations of text and graph," in *Proc. International Conf. on World Wide Web Companion. International World Wide Web Conf. Steering Committee*, 2017.

[14] L. Zhiyuan, S. Maosong, and L. Yankai, "Knowledge representation learning: A review," *Journal of Computer Research and Development*, 2016.

[15] E. Palumbo, G. Rizzo, and R. Troncy, "Knowledge graph embeddings with node2vec for item recommendation," in *Proc. European Semantic Web Conf.*, 2018.

[16] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deep walk: Online learning of social representations," in *Proc. ACM SIGKDD International Conf. on Knowledge Discovery and Data Mining*, pp. 701-710, 2014.

[17] A. Grover and J. Leskovec, "Node2vec: Scalable feature learning for networks," in *Proc. ACM SIGKDD International Conf. on Knowledge Discovery and Data Mining*, pp. 855-864, 2016.

[18] J. Li, H. Dani, and X. Hu, *Attributed Network Embedding for Learning in a Dynamic Environment*, 2017.

[19] J. Sun, Y. Xiong, and Y. Zhu, "Multi-source information fusion for personalized restaurant recommendation," in *Proc. International Acm Sigir Conf. on Research & Development in Information Retrieval*, 2015.

[20] C. Shi, B. Hu, and W. X. Zhao, "Heterogeneous information network embedding for recommendation," *IEEE Transactions on Knowledge & Data Engineering*, p. 99, 2017.

[21] J. Zhang, "Research collaboration prediction and recommendation based on network embedding in co - authorship networks," in *Proc. the Association for Information Science & Technology*, vol. 54, no. 1, pp. 847-849, 2017.

[22] X. Rong, "Word2vec parameter learning explained," *Computer Science*, 2014.

[23] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *Proc. International Conf. on Machine Learning*, 2014.

[24] W. Zhang, T. Gu, and W. Sun, "Travel attractions recommendation with travel spatial-temporal knowledge graphs," in *Proc. International Conf. of Pioneering Computer Scientists, Engineers and Educators*, 2018.

Kuang Haili was born in 1995, she is a postgraduate majoring computer science. She is currently a student in the school of Guilin University of Electronic Technology. Her research mainly focuses on personalized recommendation and artificial intelligence.

Liang Chang received his Ph.D degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, in 2008. He is currently a professor in the School of Computer Science and Information Security, Guilin University of Electronic Technology, China. His research interests include data and knowledge engineering, formal methods, and intelligent planning. (E-mail: changl@guet.edu.cn).

Yanpeng Sun received the BEng degree in computer science from received the BEng degree in computer science from the China Jiliang University, China, in 2012. He is currently working toward the M.E degree in the School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology. His research mainly focuses on data mining and data-base. (e-mail: yanpeng_sun@yeah.net)

Chenzhong Bin received the B.Eng. and M.Eng. degrees in computer science and technology from North West Polytechnical University, China, in 2002 and 2005. He is currently pursuing the Ph.D degree with the School of Information and Communication, Guilin University of Electronic Technology. His main research interests include data mining, location based service and IOT applications. (Corresponding author, E-mail: binchenzhong@guet.edu.cn, Phone: +8607732291330, Cell phone: +8613517836009, Address: No. 1, Jinji Road, Guilin City, Guangxi Prov., China 541004).