Conversation Recommender System Based on Knowledge Graph and Time-series Feature

Xiaoyi Wang
Capital Normal University, Beijing, China
2191002082@cnu.edu.cn

Yaguang Li
Capital Normal University, Beijing, China
2181002077@cnu.edu.cn

Jie Liu
Capital Normal University, Beijing, China
liujxxxy@126.com

ABSTRACT

The conversation recommender system aims to recommend high-quality items to users through interactive conversations, which requires to be seamlessly integrated between the recommendation module and the dialog module. The knowledge graph has improved the accuracy of the dialogue recommendation system to a certain extent. However, there are still some defects that make it easy to generate more general and popular responses. For overcoming these shortcomings, this paper proposes a novel framework based on knowledge graphs and time-series features called KGTF. By serializing and modelling the dialogue content, the KGTF is able to learn the feature relationship between users and items to make more accurate recommendation. Besides, the location encoding is introduced to increase the diversity of generated responses. Experiments conducted on widely adopted benchmarks show that the proposed KGTF framework is superior to the latest KGSF method.

CCS CONCEPTS

• **Information systems**; • **Information retrieval**; • **Retrieval tasks and goals**; • **Recommender systems**;

KEYWORDS

Knowledge graph, Conversation recommender system, Time-series feature, Neural network

ACM Reference Format:

Xiaoyi Wang, Yaguang Li, and Jie Liu. 2021. Conversation Recommender System Based on Knowledge Graph and Time-series Feature. In *The 5th International Conference on Computer Science and Application Engineering (CSAE 2021), October 19–21, 2021, Sanya, China.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3487075.3487149

1 INTRODUCTION

With the rapid development of the Internet, the network information data showing explosive growth, the problem of information overload has become prominent. In order to help users discover their interesting information from abundant internet data, the conversation recommender system (CRS) [1, 2] came into being, which aims to recommend high-quality items to users by interactive dialogue.

Existing studies have shown that the knowledge graph [3, 4] contains item attributes and various types of relationships, which can provide a lot of semantic information. It can improved the performance of the dialogue recommender system, but there are still some shortcomings. For example, when talking about the same item, the content of the dialogue is positive and the other is negative, the recommended results should be different. As shown in Table 1, both examples mentioned that users liked Bruce Willis, but the system gave different results due to the inconsistent order.

This paper aims to combine traditional methods with deep learning methods, and propose a timing optimizational network to model serialized dialogue data, by capturing global information and current interests to give appropriate recommendations. The experiments prove that the proposed KGTF framework has made great improvements in both recommendation and dialogue tasks. The main innovations of this article are as follows:

- Propose the TON network, which mainly uses BiGRU-attention model to serialize the dialogue content, and accurately learn the feature relationship between users and items to make appropriate recommendations.
- Since the attention mechanism does not retain sequence position information, it is easy to generate repeated words. This paper proposes to add position encoding to absorb position information to reduce the occurrence of repeated words in the generated response and increase diversity.

2 RELATED WORK

The recommendation system mainly utilizes the user's behavioral information on items to discover their individual needs, and provides users with information that meets their requirements actively through the user's interest model. Traditional recommendation system technologies include content-based recommendation, collaborative filtering recommendation, and knowledge-based recommendation.

The overall idea of the dialogue recommendation system is to support task-oriented multi-round dialogue with users, which mainly relies on the user's behavioral information on the item and the user's own expressed preferences to meet the user's interest. However, due to the rich variety of items, the user's limited learning ability leads to sparse information about the user's behavior on the item. Surprisingly, the knowledge graph contains item attributes and various types of relationships, which can provide rich item semantic information for the recommender algorithm. For example, Li et al [2]. collected a dialogue dataset focused on movie recommendations, Chen et al. proposed a KBRD [3] model for recommender

CSAE 2021, October 19–21, 2021, Sanya, China

ACM ISBN 978-1-4503-8985-3/21/10…$15.00
https://doi.org/10.1145/3487075.3487149

Table 1: Two Examples of the Outputs of Conversation Recommender Systems about Movies

Role	Example 1	Example 2
User	I am looking for a movie.	I am looking for a movie.
Recommender	The Sixth Sense is quite popular. Maybe you will like it.	Which actor do you like?
User	This movie sounds scary, so I don't want to watch it.	I like a lot, such as Bruce Willis and Henry Thomas.
Recommender	Why don't you look at Pulp Fiction, a funny comedy.	The Sixth Sense might be good for you, a horror movie.
User	That sounds good, which actor starred it?	This movie sounds scary.
Recommender	Both of these are Bruce Willis.	It played by Bruce Willis. Try it out.
User	Great! I like Bruce Willis, and I will look at both. Thank you!	Great! Thank you!

and dialogue system, Zhou et al. proposed KGSF [4], adopts KG-enhanced semantic fusion to solve the integration of dialogue and recommendation system.

The current research has greatly improved the accuracy of the dialogue recommender system based on the knowledge graph, but the existing research does not consider the timing issue [5]. On the basis of previous studies, we designed a novel method that combines KGs and time-series features.

3 MODELS

In this section, we introduce neural networks models used in the proposed model.

3.1 GRU

GRU (Gate Recurrent Unit) is a type of Recurrent Neural Network (RNN). It consists of two gates: update gate and reset gate. It can solve problems such as long-term memory and gradients in back propagation. The output value of the current node is jointly determined by the current input and the output of the previous node, it can fully learn the information before and after the dialogue content. However, it can only use historical information to make judgments on current information, and cannot use future information. Therefore, this paper considers using a two-way GRU.

3.2 BiGRU

BiGRU is a neural network model composed of unidirectional and opposite GRUs. At each moment, the input will provide two GRUs in opposite directions, the output is up to them. Therefore, the output at the current moment can be linked to the states at the previous moment and next moment, which is more conducive to the integration of dialogue history information.

3.3 Attention Mechanism

The Attention mechanism [6] assigns sufficient attention to key information to highlight local major information. Through the distribution of probability weights, the weights of word vectors at different moments are calculated, so that some words can get more attention, thereby improving the quality of feature extraction. The basic structure is shown in Figure 1

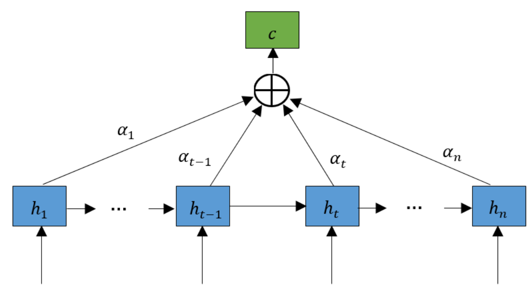

Figure 1: Attention Model.

The vector of the new hidden layer state is the cumulative sum of the product of the weight coefficient of each current hidden layer state and the initially input hidden layer state. The attention model can realize the conversion from the initial input state to the new attention state, and the information in the input sequence will greatly affect the final generated sequence. Therefore, this article uses the attention mechanism to capture global information and current interests, further highlight the key information of the dialogue content, and improve the quality of dialogue information extraction.

4 APPROACH

4.1 Encoding Knowledge Graphs

So as to fully understand user's preferences, we use two independent KGs in the dialogue and recommender systems to enhance semantic meaning. ConceptNet [7] as a word-oriented KG, and DBpedia [8] as a item-oriented KG, both of them store semantic info in the form of triples. Graph Convolutional Neural Network(GCN) uses edges to aggregate node information to generate new node representations. Unlike ConceptNet, it is necessary to consider the relationship of items. For example, when users talk about actors they like, recommenders should provide movies that are closely related them. Therefore, we choose R-GCN [9] to extract the item indication on the subgraph, and consider the type and direction of the edge to process the multi-relational data features in the knowledge base.

4.2 KG Fusion via MIM

After obtaining the node representations of word and item, the Mutual Information Maximization technique(MIM) is used to mutually enhance the data representation of the paired signals. MIM [10] can be regarded as an enhanced version of correlation, focusing on extracting temporal consistency features. Mutual information measures the degree of correlation between two random variables

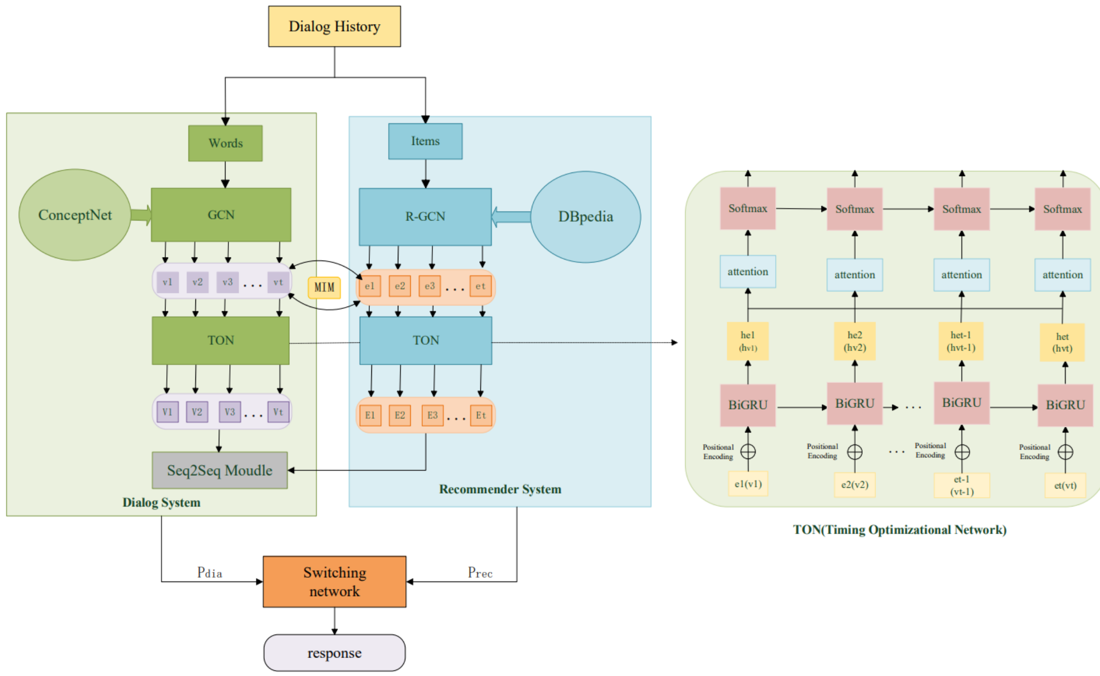

Figure 2: The Overview of Our Proposed KGTF Framework.

X and Y, and the formula is as follows:

$$MI(X,Y) = D_{KL}(P(X,Y)||P(X)P(Y)) \tag{1}$$

$$MI(X,Y) \geq E_P[g(x,y)] - E_N[g(x',y')] \tag{2}$$

where D_{KL} is the Kullback–Leibler divergence, describes the difference between two probability distributions$P(X)$and $P(Y)$, E_P and E_N represent expectations for positive samples and negative samples. MIM loss is used to pretrain the parameters of the graph convolutional network model. Treat all word-item pairs that appear simultaneously in the dialogue as positive, and the others as negative, and force the two KGs semantic spaces to be close at the beginning and try to maximize.

4.3 Timing Optimizational Network

This paper proposes a timing optimizational network(TON), which takes the representation of the fusion knowledge graph as input and passes through TON. The TON model is shown on the right in Figure 2

4.3.1 Positional Encoding. The dialogue history is time-series data, and the order between words often affects the meaning of the entire sentence. In order to avoid unnecessary misunderstandings, the words sequence needs to be considered when modeling text data, so positional encoding [11] is introduced. In this work, we use sine and cosine functions of different frequencies as follows:

$$PE(pos,\ 2i) = \sin\left(pos/10000^{2i/d_{model}}\right) \tag{3}$$

$$PE(pos,\ 2i+1) = \cos\left(pos/10000^{2i/d_{model}}\right) \tag{4}$$

where PE is a two-dimensional matrix, the row represents the word, and the column indicates the word vector; pos denotes the word position in the sentence; d_{model} represents the dimension of the word vector; i means the position of the word vector. Therefore, the above formula indicates that the sin variable and the cos variable are added to the even-numbered position of the word vector of each word to fill the entire PE matrix, and finally added to the input embedding, which completes the introduction of position encoding. In this way, the word position information is introduced, thereby reducing the occurrence of repeated words in the generated replies, and improving the diversity of the reply utterance.

4.3.2 BiGRU-attention. After the word and item data representations are respectively given relative position information, they are directly received and processed by the BiGRU layer, which can more fully understand the relationship between contexts and perform semantic coding. At the same time, the attention mechanism is introduced to assign corresponding probability weights to different word vectors, and further extract features to highlight the main information of the dialogue content. The calculation of item in the recommendation system is the same. Since the BiGRU model is regarded as two GRUs in opposite directions, the formula is simplified to equation 5). The word vector of the t-th word of the j-th sentence input at the i-th time in the dialogue system is v_{ijt}, and the specific calculation formula is:

$$h_{ijt} = BiGRU\left(v_{ijt}\right) \tag{5}$$

$$u_{ijt} = tanh\left(w_w h_{ijt} + b_w\right) \tag{6}$$

$$\alpha_{ijt} = \frac{exp\left(u_{ijt}^{T}u_{w}\right)}{\sum_{t} exp\left(u_{ijt}^{T}u_{w}\right)} \tag{7}$$

$$c_{ijt} = \sum_{i=1}^{n} \alpha_{ijt} h_{ijt} \tag{8}$$

where h_{ijt} denotes the output vector of the previous layer; w_w is the weight coefficient; u_w is the randomly initialized attention matrix. The Attention mechanism matrix is the cumulative sum of the product of the different probability weights and the state of each hidden layer, which is obtained by the normalization operation of the softmax function.

In summary, the use of TON endows the dialogue content with sequential characteristics, and the dialogue data is automatically weighted and changed to highlight key words, so that the recommendation system shows better performance.

4.4 KG-enhanced Dialog Recommender System

After obtaining the representation of the role of time-series features, in the dialogue module, transformer is used as the encoder-decoder architecture. The decoding stage uses the attention layer based on KG to fuse the information of the two knowledge graphs, and uses the recommended items, related entities and keywords to predict the next output. In the recommender module, the output is passed through a linear layer to obtain user preferences, and then the probability of each item being recommended to the user is calculated, so as to make ranking recommendations for the items.

Until now, we have built an end-to-end framework to connect the recommendation module and the dialogue module to achieve a friendly integration between the systems.

5 EXPERIMENTS

In this section, we introduce the details of our experiments, including dataset, implement as well as results.

5.1 Dataset

The experiment selects REcommendations through DIALog (REDIAL) as the dataset. The number of conversations and utterances are 10006 and 182150.Moreover, the total number of users and movies are 956 and 51699 respectively. And the training set, validation set and test set are divided according to the ratio of 8:1:1. In addition, we also introduce related entities and relationships from ConcepNet and DBpedia.

5.2 Implement

The model is implied by the pytorch deep learning framework and Python programming language. And the experiment running environment is JetBrains PyCharm software, ubuntu20.04 system, etc. The embedding dimensions of the dialogue system and the recommendation system are set to 300 and 128 respectively; the number of layers L of GCN and R-GCN are both 1. During the training, we set the batch size is 32, the learning rate is set to 0.001 and the gradient is limited to [0,0.1].

5.3 Evaluation Metrics

We use two different indicators to evaluate the dialogue and recommendation modules respectively. Dialogue evaluation includes automatic and human evaluation. We use the Distinct indicator to determine whether there are a ton of universal and repetitive responses. The larger the Distinct-n, the higher the variety of responses generated. At the same time, in the manual evaluation, since the responses we hope to generate are practical suggestions related to the item, we invited 5 commenters to rate the generated responses on fluency and information volume. We selected 100 rounds of dialogues from the test set for evaluation with a score range of 0-3, and finally used the average score of 5 commenters for comparison.

The evaluation metric of the recommended module is recall@k. It refers to the ratio of the first k results returned from the last recommendation list sorted by score, and then the number of relevant results retrieved from it to the number of all relevant results in the library.

5.4 Compared Methods

We compared KGTF with some mainstream dialogue recommendation systems:

- TextCNN: It proposes a CNN-based model to complete sentence-level classification tasks by extracting features.
- Transformer: It uses a Transformer-based framework to get suitable responses for dialog module.
- REDIAL[2]: It is mainly composed of dialogue generation system based on HRED, a recommender system based on auto-encoder.
- KBRD[3]: It proposes a framework named KBRD and it connectes the recommendation and the dialogue system through the dissemination of knowledge.
- KGSF[4]: It proposes a KG-based semantic fusion approach, utilizes MIM to align the semantic spaces between words and items by using two external KGs.

Among the above baseline models, TextCNN is a recommended method, and Transformer [6] is the most advanced text generation method. "REDIAL", "KBRD" and "KGSF" are recommended methods for dialogue. We named the newly proposed model KGTF.

5.5 Results

In this section, we will introduce the experimental results, Table 2 shows the performance of the different methods.

5.5.1 Recommended Module Performance. In order to evaluate the effectiveness of the recommendation system, we evaluated Recall@K. The results in Table 2 show that the proposed model has achieved a certain improvement in the evaluation of Recall@1, 10 and 50.

In addition, this study also used ablation studies (Figure 3), where KGTF(B) stands for only utilizing the BiGRU Model, KGTF(A) stands for only adopting Attention mechanism and KGTF(-P) means using the BiGRU-Attention model without positional encoding. It clearly shows that the time-series features and location information is valid. Overall, the improvement of all three indicators proves the efficiency of our model.

Table 2: Results on Recommendation System

Model	Recall@1	Recall@10	Recall@50
TextCNN	0.013	0.068	0.192
ReDial	0.021	0.140	0.311
KBRD	0.031	0.151	0.336
KGSF	0.034	0.179	0.376
KGTF	0.036*	0.196*	0.388*

Note: The error is less than 0.05.

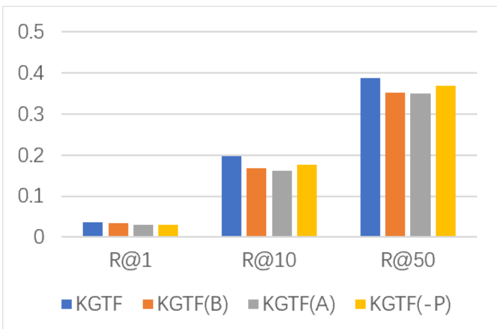

Figure 3: Ablation Study on Recommender Task.

Table 3: Automatic and Human Evaluation

Model	Dist-2	Dist-3	Dist-4	Info
Transformer	0.148	0.152	0.137	1.08
ReDial	0.225	0.240	0.229	0.97
KBRD	0.268	0.368	0.423	1.18
KGSF	0.289	0.435	0.529	1.41
KGTF	0.399*	0.579*	0.687*	1.56*

Note: The error is less than 0.05.

5.5.2 Dialog Module Performance. Table 3 shows the evaluation results of the baseline model and the method we proposed in conversation generation. For automatic evaluation, we evaluated distinct@k. The results show that the proposed model is significantly better than baselines in 2-gram, 3-gram and 4-gram results. This shows that our model can produce more diverse content to greatly improve the performance of the dialogue system. Moreover, it is extremely necessary to introduce the time-series of dialogue content.

For manual evaluation, we let human annotators rate the informativeness of the conversation content. Compared with the previous best KGSF, KGTF reaches better performance by +0.15 informativeness score. It greatly shows that our model can effectively use contextual information and produce smooth and informative responses.

6 CONCLUSION AND FUTURE WORK

This paper proposes a framework for conversation recommender system based on knowledge graph and time series features, namely KGTF. Two external knowledge graphs are used to enhance the semantic representation of word and item, and mutual information maximization technology is used to align the semantic space. And use the TON model to give different recommendations according to the order of the dialogue content to improve the performance of the dialogue recommendation system. The experiments proved that KGTF produces better performance than the baseline in terms of dialogue and recommendation. In the future, we are committed to proposing better methods to further improve the performance of the dialogue recommendation system and its application in the industry.

ACKNOWLEDGMENTS

This work is supported partially by National Science and technology innovation 2030 major projects (2020AAA0109700), National Natural Science Foundation of China (62076167), and Beijing Municipal Education Commission-Beijing Natural Fund Joint Funding Project (KZ201910028039).

REFERENCES

[1] D. Jannach, A. Manzoor, W. Cai and L. Chen (2020). A survey on conversational recommender systems, arXiv:2004.00646, [online] Available: http://arxiv.org/abs/2004.00646.
[2] Raymond Li, Samira Ebrahimi Kahou, and Hannes Schulz (2018). Towards Deep Conversational Recommendations. 2018, 9748–9758.
[3] Qibin Chen, Junyang Lin, and Yichang Zhang (2019). Towards Knowledge-Based Recommender Dialog System. 2019, 1803–1813.
[4] Kun Zhou, Wayne Xin Zhao, and Shuqing Bia (2020). Improving Conversational Recommender Systems via Knowledge Graph based Semantic Fusion. In KDD 2020.
[5] Liu J, Yang Y, Lv S, Wang J, Chen H (2019). Attention-based BiGRUCNN for chinese question classifcation. J Ambient Intell Hum Comput, pp 1-12.
[6] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin (2017). Attention is All you Need, pp. 5998–6008.
[7] Robyn Speer, Joshua Chin and Catherine Havasi (2017). ConceptNet 5.5: An Open Multilingual Graph of General Knowledge. In AAAI 2017. 4444–4451.
[8] BIZER C, LEHMANN J, KOBILAROV G, *et al.* (2009). DBpedia-a crystallization point for the Web of data[J]. Web Semantics Science Services & Agents on the World Wide Web, 7(3), 154-165.
[9] Michael Edwards and Xianghua Xie (2016). Graph Convolutional Neural Network. In BMVC 2016.
[10] Yi-Ting Yeh and Yun-Nung Chen (2019). Qainfomax: Learning robust question answering system by mutual information maximization. In EMNLP-IJCNLP. 3361–3366.
[11] Y. Yang, Y. Tong, S. Ma and Z.-H. Deng (2016). A position encoding convolutional neural network based on dependency tree for relation classification. In EMNLP 2016. 65-74.