Multi-Information Preprocessing Event Extraction With BiLSTM-CRF Attention for Academic Knowledge Graph Construction

Chao Chang, Yong Tang, Yongxu Long, Kun Hu, Ying Li, Jianguo Li, and Chang-Dong Wang

***Abstract*— Academic knowledge graph is an important application of knowledge graph in the vertical field of academia. At present, the construction of the academic knowledge graph is mainly completed by extracting published academic papers, authors, publications, and other information from related databases. However, academic information is not just information of published papers. Scholars' academic activities include participation in academic conferences, visiting and making presentation, and so on. However, the above academic information is hidden in natural language texts and cannot be directly stored in academic knowledge graph. This article proposes an approach named construct-SCHOLAT knowledge graph to construct an academic event knowledge graph based on academic social network SCHOLAT. The construction framework mainly consists of two parts: data preprocessing and event extraction. In the data preprocessing, we propose a knowledge graph embedding method to represent scholars' academic social feature. In the event extraction, we concatenate the preprocessed scholar vector with academic we-media blog text into the extraction model based on BiLSTM-CRF fused with attention mechanism. The extracted events are added to academic knowledge graph, and a public relationship exists between the event and the scholar. Compared to the previous methods, our framework has an excellent performance after experimental verification. To the best of our knowledge, this is the first study to use the scholar academic social information of the scholar who edited the text as the event extraction input information. In addition, we publish a Chinese event extraction dataset SCHOLAT academic event extraction.[1] The dataset includes academic we-media blog and the social behavior embedding vector of the scholar. All the data in this dataset are derived from the academic social network SCHOLAT.**

***Index Terms*— Academic knowledge graph, event extraction, knowledge graph embedding, SCHOLAT.**

Manuscript received 23 September 2021; revised 24 January 2022 and 18 April 2022; accepted 4 June 2022. This work was supported by the National Natural Science Foundation of China under Grant U1811263, Grant 62077045, and Grant 61876193. *(Corresponding author: Yong Tang.)*

Chao Chang, Yong Tang, Yongxu Long, Kun Hu, Ying Li, and Jianguo Li are with the School of Computer Science, South China Normal University, Guangzhou, Guangdong 510632, China (e-mail: changchao@m.scnu.edu.cn; ytang@m.scnu.edu.cn; longyongxu@m.scun.edu.cn; lkome@m.scun.edu.cn; 2345903@qq.com; jianguoli@m.scnu.edu.cn).

Chang-Dong Wang is with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510275, China (e-mail: changdongwang@hotmail.com).

Digital Object Identifier 10.1109/TCSS.2022.3183685

[1]https://www.scholat.com/research/opendata

I. Introduction

SINCE knowledge graph was proposed by Google in 2012, it has made significant progress and development. Massive Internet information is transformed into more intelligent and structured knowledge. As the cornerstone of artificial intelligence (AI), knowledge graph has excellent performance in search, recommendation, and other fields [1]. In the field of academic knowledge graph, academic information, such as scholars, papers, and coauthor relationships, is stored in academic knowledge graph. At present, the research on academic knowledge graph focuses more on research fields, academic papers, coauthoring relationships, and research institute while ignoring other academic activities of scholars. When we want to analyze the academic information of a scholar, the above information is far from enough. Scholar's academic activities, such as making presentation, participating in academic conferences, and getting a promotion, play an important role in analyzing scholars' influence, recommendation, search, and other application fields. Therefore, how to integrate more academic activities of scholars into academic knowledge graph is a problem to be solved.

Events are at the core of human society. The evolutionary laws and patterns of events occurring one after another in the time dimension are a kind of very valuable knowledge [2]. In recent years, event knowledge graph has drawn a growing number of scholars' attention. In [3], an approach of extracting events from news articles to create event-centric knowledge graphs (ECKGs) is presented. Through ECKGS, the reconstruction of news storylines is promoted. In [4], a general method that leverages the Wikidata event knowledge base to produce semantic annotations of news articles is proposed. Filter for properties belonging to specific event patterns that are automatically inferred is provided. In [5], recurrent event network (RE-NET), a novel autoregressive architecture for predicting future interactions, is developed. Future facts could be inferred in a sequential manner based on RE-NET. In [6], a temporal graph learning method with heterogeneous data fusion for predicting concurrent events of multiple types and inferring multiple candidate actors simultaneously is studied. Costa *et al.* [7] presented the Event-QA dataset for answering event-centric questions over knowledge graphs. In [8], a novel method representing social events and their relationships as a knowledge graph is proposed for decomposing and discovering social events. In addition,

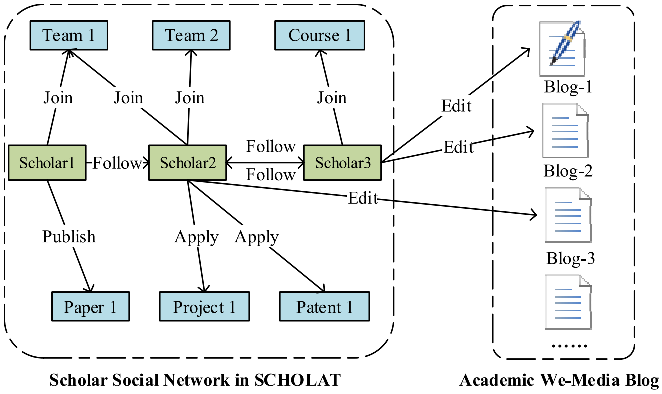


Fig. 1. Data source for constructing the academic knowledge graph is based on SCHOLAT.

event knowledge graph has also been applied to the fields of biomedicine [9], finance [10], tourism [11], and so on. However, event knowledge graph has not been researched and applied deeply in the field of academic knowledge graph.

Extracting events from academic activity information and importing events into academic knowledge graph are an innovative application. In this article, we propose to construct an academic knowledge graph named SCHOLAT knowledge graph (SKG) that contains scholars, scholars' social relationships, papers, projects, patents, academic event, and so on. SCHOLAT[2] is implemented as a scholar-oriented social network to form an academic community to let scholars establish connection with other researchers [12]. As shown in Fig. 1, the information used to construct SKG in SCHOLAT includes structured data and unstructured data. Structured data include scholars, teams, courses, papers, projects, patents, and related academic social relationships. Unstructured data are academic we-media blog edited by scholars on the scholar social network SCHOLAT. This article addresses the issue of Chinese text event extraction.

For structured data, we directly import them into SKG to construct scholar entities, paper entities, project entities, patents entities, team entities, and course entities. For unstructured data, we extract events from academic we-media blogs and add the events as entities to SKG. Due to the different writing habits, academic, and social attributes of scholars in academic social network SCHOLAT, we propose a novel event extraction model. Different from the previous event extraction method that only focuses on the text itself, we comprehensively consider the social features of the text author, namely, scholar and the text. Fig. 2 shows several examples of event extraction from academic we-media blog sentence edited by two scholars in SCHOLAT. The academic we-media blog edited by Scholar "Yong Tang" and "Chao Chang" are different in content and wording. However, from their blog texts, we could extract the same type of academic event that a conference has been held. Therefore, we try to extract the academic social attributes of the scholars who edit the text as a continuous vector and use it for event extraction.

[2]www.scholat.com

In order to extract academic events from academic we-media texts to construct an academic knowledge graph SKG, we propose a novel event extraction model called multi-information preprocessing BiLSTM-CRF attention (MIP BiLSTM-CRF attention), which is a multi-information preprocessing event extraction method. In the information preprocessing phase, academic social attributes of the scholar are presented as a continuous vector representation based on multilayer convolutional neural network with input vector reshaping. Meanwhile, we input the academic we-media text sentence by sentence into Bert [13] to complete the text preprocessing. In the event extraction phase, we concatenate the scholar embedding vector and text embedding vector into BiLSTM-CRF fused with attention mechanism to complete event extraction. This makes the representation of every text character have the academic social features of the scholar who edited it.

In summary, the contributions of this article are summarized as follows.

1) This article for the first time studies the extraction of scholars' social relations and scholars' academic activity events to construct an academic knowledge graph.
2) Event extraction is the core step of the academic knowledge graph construction. We propose a novel event extraction model, MIP BiLSTM-CRF attention, which is to concatenate the preprocessed scholar vector and text vector into BiLSTM-CRF fused with attention mechanism to extract academic event. This concatenation enables every character in the sentence to have the academic social attributes of the scholar who edited this sentence.
3) We publish a Chinese event extraction dataset SCHOLAT academic event extraction (SAEE). The SAEE dataset consists of two parts: the vector representation of scholars and the academic we-media blog text edited by scholars.

This article is organized as follows. We review related work in Section II and provide preliminaries in Section III. Then, we illustrate our framework and algorithm in Section IV. We report the experimental results in Section V. Finally, we conclude this article in Section VI.

II. Related Work

A. Academic Knowledge Graph

Research on the knowledge graph in the academic field is promoted. The Microsoft Academic Knowledge Graph [14] is composed of their AI agents extraction information. It contains information about scientific publications and related entities, such as authors, institutions, journals, and research fields. Science Knowledge Graph [15] is a knowledge graph designed for scientific purpose, consisting of concepts, experts, and papers. The concepts and their relationships are extracted from ACM computing classification system. The Open Academic Atlas [16] is an academic knowledge graph, which connects two graphs: Microsoft Academic Graph and Science Knowledge Graph. Acemap Knowledge Graph [17] provides an academic heterogeneous map, including a variety of academic entities and corresponding attributes. It includes papers,

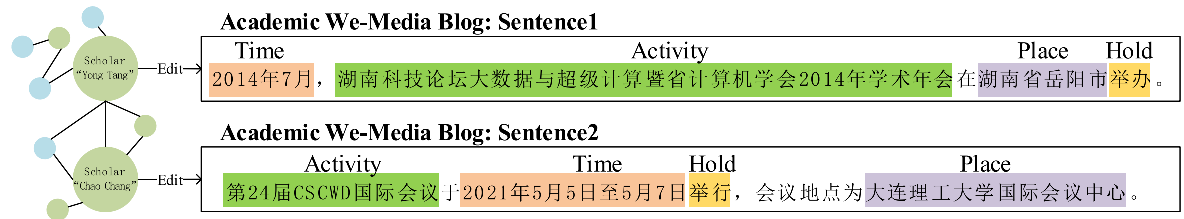


Fig. 2. Examples of event extraction from academic we-media blog sentences edited by two scholars in SCHOLAT.

authors, fields, institutions, journals, conferences, alliances, and supports authoritative and practical academic research. Web of Scholars [18] is a scholar knowledge graph in the computer field. Relying on the knowledge, it provides services for semantic querying as well as recommendations. However, the research on academic knowledge graph is limited to the relationship between scholars' research fields, papers, and relationships but do not cover the relationship between the scholar and the event in which the scholar has participated in.

B. Knowledge Graph Embedding

Knowledge graph embedding refers to embedding the triplets of the knowledge graph containing entities and relationships into a continuous vector space [19]. The typical knowledge graph embedding process usually has three steps, which are representing entities and relationships, defining score functions, and learning entity and relationship representations. In the choice of the score function, some researchers choose the translation distance models [20]–[22]. Some researchers choose the semantic matching models [23]. Some researchers choose the neural network models [24], [25].

C. Event Extraction

In order to construct an academic event knowledge graph, we need to complete the task of extracting events from natural language texts. In the early stage of the development of event extraction research, the method based on pattern matching [26], [27] is used. Due to pattern matchings' weak generalization ability and the need for a large number of manual annotations for template construction, event extraction based on machine learning is more popular. In [28], a hierarchical modular event argument extraction model is proposed for better complete event extraction by classifying different event roles. In [29], an event extraction method using bleached statements to give a model access to information contained in annotation manuals is presented. Li *et al.* [30] proposed a knowledge base-driven tree-structured long short-term memory networks framework to apply in biology.

However, related research about Chinese event detection and extraction research is still in its infancy. In [31], BILSTM is used to encode the semantics of the words in the entire sentence into sentence-level features without any syntactic analysis. Then, the convolutional neural network is used to capture the prominent local lexical features to eliminate the ambiguity of the trigger, thereby completing the Chinese event extraction. In [32], an end-to-end document-level framework is proposed for Chinese financial event extraction and published a large-scale real dataset of Chinese financial announcements. In [33], the trigger-aware lattice neural network is proposed to cope with the problem of word trigger mismatch in word-based models and the ambiguity of trigger words would affect trigger classification. However, there is no relevant research on event extraction in the academic field, and no relevant datasets have been released. In addition, there are some studies on open-domain event extraction [30], [34], [35]. At present, the research of open-domain event extraction is focused on event detection, topic detection, relationship mining, and so on.

III. Preliminary

In this section, the definitions and problem statement are introduced.

A. Definitions

Definition 1 (Knowledge Graph): A knowledge graph is defined as $G = (\varepsilon, R, T)$, where $\varepsilon = \{e_1, e_2, \ldots, e_{|\varepsilon|}\}$ denotes the set of entities and $R = \{r_1, r_2, \ldots, r_{|R|}\}$ denotes the set of relations. $T \subseteq \varepsilon \times R \times \varepsilon$ denotes the set of triplets and a triplet could be expressed as (e_s, e_r, e_o). The letters $\mathbf{e}_s, \mathbf{e}_r, \mathbf{e}_o \in \mathbb{R}^k$, respectively, denote the k-dimensional embeddings of e_s, e_r, e_o.

Definition 2 (Event): Given a scholar-centered social network SCHOLAT, an event is defined as $E = (\text{Sid}, Ti, Ty, \text{Arg})$, where $\text{Sid} \in \varepsilon$ denotes the id of the scholar who edited the event, Ti denotes the time the event was posted on the website, $Ty = \{t_1, t_2, \ldots, t_{|Ty|}\}$ denotes the event type, and $\text{Arg} = \{a_1, a_2, \ldots, a_{|\text{Arg}|}\}$ denotes the event argument and its corresponding value.

Definition 3 (Event Knowledge Graph): An event knowledge graph is defined as $\text{EG} = \{\varepsilon, E, R, T\}$, where $\varepsilon = \{e_1, e_2, \ldots, e_{|\varepsilon|}\}$ denotes the set of entities, $E = \{ev_1, ev_2, \ldots, ev_{|E|}\}$ denotes the set of events, $R = \{r_1, r_2, \ldots, r_{|R|}\}$ denotes the set of relations, and $T \in (\varepsilon \times R \times \varepsilon) \cup (\varepsilon \times R \times ev)$ denotes the relationship between entity and entity or entity and event.

B. Problem Statement

Definition 4 (Knowledge Graph Embedding): Given a knowledge graph $G = (\varepsilon, R, T)$, knowledge graph embedding is to embed components of G such as entities and relations into continuous vector spaces $e_s = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$.

Definition 5 (Event Extraction): Given an input sentence consisting of k words $\{\text{ch}[1], \text{ch}[2], \ldots, \text{ch}[k]\}$, the task is to identify event types and event arguments in the text.

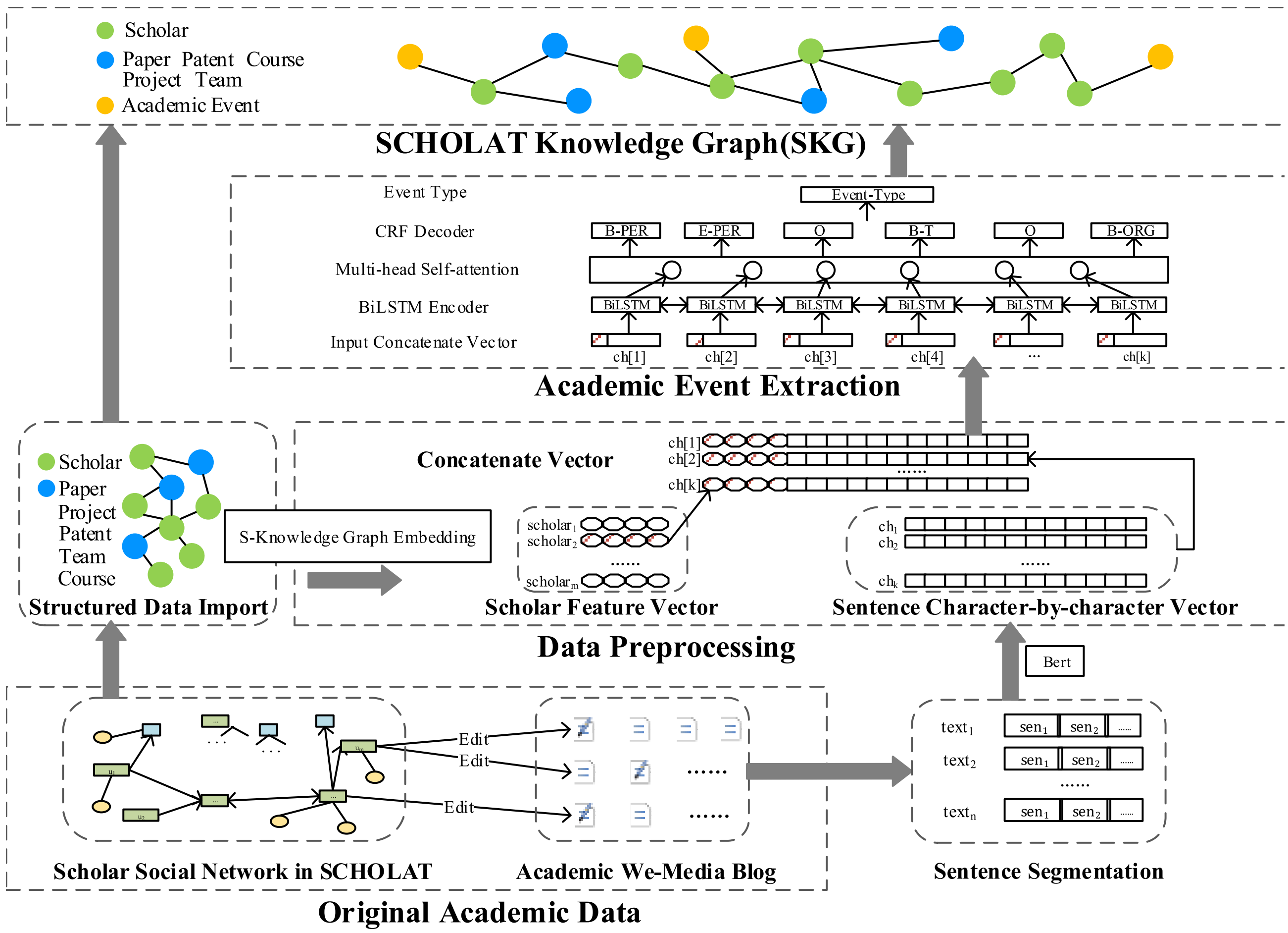


Fig. 3. Overall framework of C-SKG.

Definition 6 (Academic Knowledge Graph Construction): Using the top-down construction pattern, the task is to extract scholars, courses, teams, academic events as entities, and their relationships from academic social network SCHOLAT to build an academic knowledge graph.

The problem we solve in this article is to extract academic events and event arguments from academic we-media texts and use them with other scholar social information to construct an academic knowledge graph. There are two issues associated with the problem abovementioned. The first issue is knowledge graph embedding in the event extraction data preprocessing phase, which solves the problem of scholars' academic social attributes vector representation. The second issue is event extraction, which solves the problem of extracting effective academic event information from natural language texts.

IV. Proposed Method

A. Framework Overview

The core idea of construct-SCHOLAT knowledge graph (C-SKG) is to build an event extraction model for constructing a knowledge graph of scholar social events. SKG includes two layers. The bottom layer is the original knowledge graph of scholar social network, and the upper layer is the academic event layer. The two layers are connected by the relationship between scholar and academic event.

As shown in Fig. 3, the C-SKG model framework and information flow are illustrated. In the scholar-centered knowledge graph SKG, we embed each scholar's knowledge graph to express the academic social attribute of scholars. For the part of the unstructured academic we-media blog, we segment it and preprocess it. Then, we concatenate the scholar feature vector and the character segmentation vector into the event extract model. Finally, the extracted events are added to the SKG to construct knowledge graph of SCHOLAT academic event. The framework mainly consists of two parts: data preprocessing and academic event extraction.

1) *Data Preprocessing:* There are two main parts of data preprocessing, namely, S-Knowledge graph embedding and Chinese sentence character-by-character vector representation. As we know, knowledge graph embedding is an excellent method to express entity, relationship, attribute, and other information into vector space [36]. Using knowledge graphs for link prediction is a typical application of knowledge graph embedding [37]. However, the method of link prediction could also help knowledge graph embedding to get better parameters. In this part, we propose a knowledge graph embedding method based on link prediction to represent the features of each scholar entity in SKG. Expressive features include scholars' personal attributes, academic social

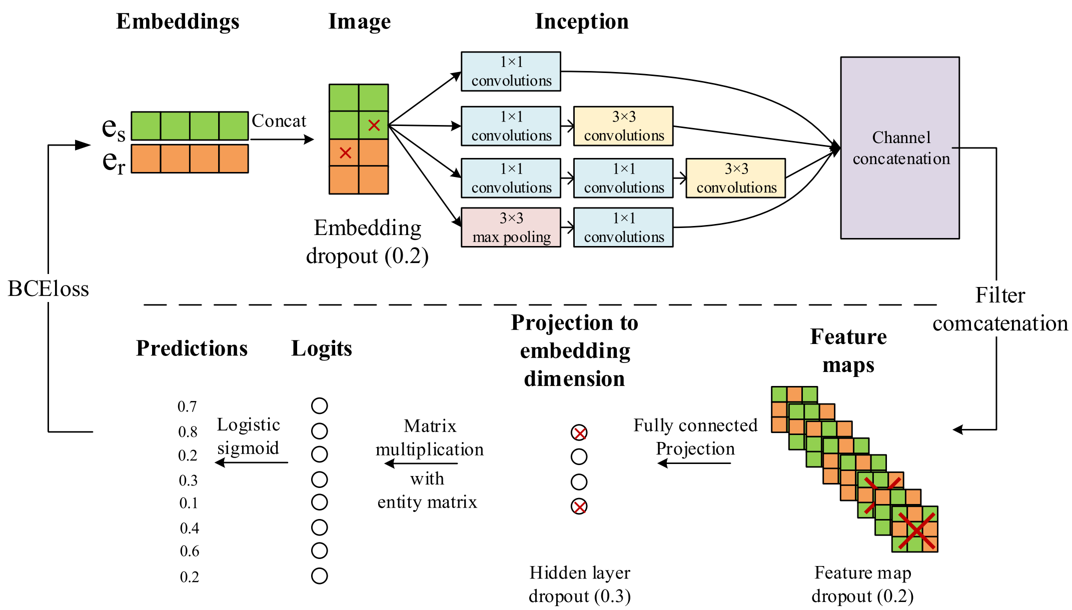


Fig. 4. Knowledge graph embedding link prediction scoring function model.

relationships, academic achievement, and so on. Bert is used to complete sentence character-by-character vector representation.

2) *Academic Event Extraction:* In this part, we concatenate the preprocessed scholar vector with the text character vector edited by the scholar. Then, the concatenated vector is input to BiLSTM with multihead self-attention layer. Finally, academic event arguments and event types are extracted. The extracted events are linked to the scholars who posted the events. Ultimately, we achieve the goal of building the knowledge graph of scholar social events (SKG).

In summary, the C-SKG framework proposed in this article could extract the behavioral feature of scholars in SCHOLAT and could be used to construct an academic event knowledge graph in the social network of scholars. In this section, we describe the C-SKG model considering the social attributes in detail. The model is mainly composed of two parts, namely, data preprocessing and academic event extraction.

B. Data Preprocessing

1) Academic We-Media Text Processing: For the input sentence $\{\mathrm{ch}[1], \mathrm{ch}[2], \ldots, \mathrm{ch}[k]\}$, Bert [13] is used for sentence preprocessing. Bert is a language representation model proposed by Google, which mainly includes two processes: pretraining and fine-tuning. The input word vector of Bert is the sum of the three vectors of token embedding, segment embedding, and position embedding. In the pretraining part, two unsupervised tasks, masked LM and next sentence prediction, are used to train the model. In the fine-tuning part, the input and output of a specific task are inserted into the Bert. Then, all parameters are fine-tuned end-to-end. In this article, we use Bert-base-Chinese as the pretraining model. After processing, the Chinese characters in the sentence are represented as continuous vectors $\mathrm{ch}[t] = \{\alpha_1, \alpha_2, \ldots, \alpha_m\}$.

2) Knowledge Graph Embedding: In order to express the scholar attributes and academic social attributes in SCHOLAT through a continuous vector space, we choose to embed the scholar entities in SKG. The 2-D convolutional neural network enhances the expressive force of the model through the interaction between embedding, which can extract more feature interactions in a shorter time. The idea is to learn the knowledge graph embedding with link prediction as the optimization goal. The overall algorithm of knowledge graph embedding link prediction scoring function is shown in Fig. 4. We first make an initial embedding of e_s, e_r in T. Then, we put the vector into a convolutional neural network for training. Through backpropagation, the parameters are updated to achieve more accurate embedding purposes.

The scoring function is defined as follows:

$$\varphi_r(e_s, e_o) = f(\mathrm{vec}(f([\overline{e_s}; \overline{e_r}] \oplus \omega))W)e_o \tag{1}$$

where $\overline{e}_s$ and $\overline{e}_r$ denote vector reshaping of e_s and e_r, respectively.

We concatenate the two reshaped vevtors $\overline{e_s}$ and $\overline{e_r}$ into an "image" $\overline{e_s}$ and $\overline{e_r}$. Then, cut the "image" and batch normalization into a four-layer inception convolutional layer. $\oplus$ denotes the convolution layer. ω denotes the filters. After convolution, we use rectified linear unit as the activation function. Then, enter the output layer of the previous step into the channel concatenation layer. After filtering concatenation, the feature maps are obtained. Through full link operation, the hidden layer is obtained. Then, we get a vector whose dimension is the total number of entities after matrix calculation. In order to prevent overfitting, we use dropout in some steps in the model. Finally, we use the activation function sigmoid to map the logistics layer to $(0, 1)$, as shown in the following equation:

$$\mathrm{sigmoid}(x) = \frac{1}{1 + e^{-x}}. \tag{2}$$

In the backpropagation part, BCEloss is used as a loss function to optimize the parameters in e_s and e_r as follows:

$$\text{BCEloss}(X_i, y_i) = -\omega_i\left[y_i \log \mathrm{x}_i + (1 - y_i)\log(1 - x_i)\right] \quad (3)$$

where $X_i \in \mathbb{R}^{1\times|\varepsilon|}$ is a score vector predicted by the model and y_i is the label vector processed by label smoothing with demension $\mathbb{R}^{|\varepsilon|\times 1}$.

C. Event Extraction Model

In the event extraction part, we use a unified event extraction model. In this model, we finish the event-type recognition and event metadata extraction. The overall algorithm of event extraction is shown in Fig. 3.

The preprocessed scholar vector and the text vector are concatenated into the event extraction model. The concatenation setting makes each sentence not only have the meaning of we-media blog text but also have the academic social attributes of the scholar who edited the text.

LSTM is a variant of recurrent neural network, which is used to process time series data. LSTM selects the addition or deletion of the cell state through the gate structure. The memory unit has three gate structures: input gates, forget gates, and output gates, which are used to maintain and update the cell state. Here, we use i_t, f_t, o_t, and C_t to represent the three kinds of structures and nerve cell states corresponding to time t.

The first step of LSTM is to delete some information in nerve cells, which is determined by the sigmoid layer of the forget gate. The input of the current step is x_t and the hidden layer output h_{t-1} from the previous layer. The calculation formula is shown in the following equation:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right). \quad (4)$$

The second step is to determine which new information should be stored in the cell state. The input gate determines the value to be updated by the sigmoid layer, and the new candidate value $\widetilde{C}_t$ is created by the tanh layer and added to the nerve cell. The calculation formula is shown in the following equations:

$$i_t = \sigma\left(W_t \cdot \left[h_{t-1}, x_t\right] + b_i\right) \quad (5)$$

$$\widetilde{C}_t = \tanh\left(W_c \cdot \left[h_{t-1}, x_t\right] + b_c\right). \quad (6)$$

After the state of the entire nerve cell is updated, we multiply the previous state of the nerve cell C_{t-1} by f_t and delete the information that is determined to be useless. Then, we add $i_t \cdot \widetilde{C}_t$ after it to get the current update value of the nerve cell state. The calculation formula is shown in the following equation:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t. \quad (7)$$

Finally, we get the output of the information through the output gate. The output h of the LSTM unit is as follows:

$$h_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \cdot \tanh(C_t). \quad (8)$$

In order to better capture the context information of the sentence and obtain the global sequence information of the sentence more comprehensively, we use BILSTM to solve the problem that the LSTM cannot process the subsequent information. BILSTM contains a forward and backward LSTM. We connect the forward hidden state $\overrightarrow{h_t}$ and the backward hidden state $\overleftarrow{h_t}$. Then, the hidden vector h_t was obtained as follows:

$$h_t = \overrightarrow{h_t}\,||\,\overleftarrow{h_t}. \quad (9)$$

The BILSTM output vector is $H = \{h_1, h_2, \ldots, h_k\}$. Because each character has a different contribution in the event extraction task, we introduce a multihead self-attention [38] layer after the BiLSTM layer. Then, we calculate the self-attention of each part as shown in the following equation:

$$\text{head}_i = \varpi(QK^T)V. \quad (10)$$

Then, we concatenate all the combinations learned by head as follows:

$$\text{MultiHead} = \text{concat}\left\{\text{head}_1, \text{head}_2, \ldots, \text{head}_p\right\}W^o. \quad (11)$$

Finally, we input the vector to the CRF layer. Through the CRF decoder, the argument of the event is identified. According to the results of event argument identification, we complete event identification and classification.

For clarity, the entire academic event extraction algorithm is summarized in Algorithm 1.

Algorithm 1 MIP BiLSTM-CRF Attention Academic Event Extraction Model

Input: The set of scholar academic social network, $G = (\varepsilon, R, T)$; The set of academic we-media blog texts, T;
Output: Academic event, $E = (\text{Sid}, Ti, Ty, Arg)$;
1: Embedding scholar academic social network information, obtain scholar feature vector;
2: Use Bert to complete text-by-sentence vector representation;
3: Concatenate the feature vector of scholar Sid with the character vector in the corresponding text Ti;
4: Input the concatenated vector into the BiLSTM layer;
5: Input the output of the previous layer into the muli-head self-attention layer;
6: Input the output of the previous layer into the CRF layer;
7: Obtain recognized academic event argument Arg and event type Ty;
8: **return** E;

V. Experimental Evaluation

In this section, the experimental dataset is introduced, and the experimental results are analyzed.

A. Experimental Settings

In order to evaluate the academic feature representation model of scholars in the data preprocessing stage, which is used to complete the knowledge graph embedding task, we have verified it on three datasets. Table I shows the details of entities, relationships, and triplets in each dataset.

1) WN18RR [39] is a subset of WN18. To remove the WN18 reverse relations, WN18RR is created by keeping

TABLE I
DETAILS OF THE DATASETS USED TO KNOWLEDGE GRAPH EMBEDDING

Dataset	| E |	| R |	#Triples		
			Train	Valid	Test
WN18RR	40,943	11	86,835	3,034	3,134
FB15k-237	14,541	237	272,115	17,535	20,046
YAGO3-10-DR	122,837	36	732,556	3,390	3,359

just one relation from each pair of reverse relations. The resulting dataset WN18RR has 40 943 entities in 11 relations. Most of the triple relationships are composed of hyponym and hypernym relations.

2) FB15k-237 [39] is a series of triplets extracted from FB15k by removing inverse and near-duplicate relations. The dataset FB15k-237 contains 310 116 triples with 14 541 entities and 237 relations. The data topics are mainly about movies, actors, sports, and so on.
3) YAGO3-10-DR [39] is a subset of YAGO3, which mainly integrates data from three sources: Wikipedia, WordNet, and GeoNames. YAGO3-10-DR is created by removing a triple from the test and validation sets if it belongs to any symmetric relation and its entity pairs are directly linked in the training set.

1) Datasets: In order to evaluate the event extraction model, which is used to complete the knowledge graph event extraction task, we have verified it on the dataset termed SAEE published in this article.

SAEE is a dataset based on academic social network SCHOLAT for Chinese event extraction. We selected the academic behaviors and we-blogs of the top 10 000 active scholars of SCHOLAT. Then, we remove the sensitive information and related confidential information and junk information. We complete the representation of scholars through KGE and manually annotate the we-blog sentence by sentence. The SAEE contains two parts: 1) the social attribute embedding representation vector of the academic we-media blog author namely scholar and 2) the sentence-level Chinese academic blog event annotation. Table II lists the event types of SAEE. In SAEE, we have defined ten types of academic events, including publication of academic achievements, getting honors, and holding academic activities. We selected 800 blogs published by 60 scholars, annotated them, and published them on the website.

As shown in Table III, compared with ACE 2005 Chinese [40], which is commonly used for event extraction in the Chinese field, the SAEE dataset has academic social information of event publishers. The raw corpus of SAEE is published by the scholars of registered scholars of SCHOLAT, all of which are academic we-media blogs. SAEE has 60% more files than the ACE dataset.

2) Baselines: We compare our knowledge graph embedding model in the data preprocessing stage with the following baselines.

1) DisMult [23] was published in 2015. It is a framework of learning representations of entities and relations in KBs using the neural-embedding approach.
2) ConvE [24] was published in 2018. It is a multi-layer convolutional network model for knowledge graph embedding for link prediction.
3) RotaE [41] was published in 2019. It is able to model and infer various relational patterns. The model defines each relationship as a rotation from the source entity to the target entity in the complex vector space.
4) Composite neighborhood (CoNE) [42] was published in 2020. It is an end-to-end framework CoNE, utilizing composite neighbors to enhance the existing KGE methods.

We compare MIP BiLSTM-CRF attention with the following baselines.

1) JointBeam [43] was published in 2013. It is a joint framework based on structured prediction, which extracts triggers and arguments together.
2) DMCNN [44] was published in 2016. It is a dynamic multipooling convolutional neural network, which could extract lexical- and sentence-level features without using complicated NLP tools.
3) JEE [45] was published in 2016. It is a joint framework with bidirectional recurrent neural networks to extract event.
4) RBPB [46] was published in 2016. It uses trigger embedding, sentence-level embedding, and pattern features as the features for trigger classification, which considers two kinds of argument relationships.
5) JMEE [47] was published in 2018. It is a framework to jointly extract multiple event triggers and arguments by introducing syntactically shortcut arcs to enhance information flow and attention-based graph.
6) Zero-shot transfer learning [48] was published in 2018. It is a transferable neural architecture. Using the zero-shot learning model to extract invisible types from existing types and transfer knowledge.
7) Joint transition [49] was published in 2019. It is a transition-based model for jointly predicting nested entities, event triggers, and their semantic roles in event extraction. The model detects entities and event mentions by using an incremental left to right reading order.
8) Cross-lingual structure transfer for zero-resource event extraction [50] was published in 2020. It is a cross-lingual structure transfer framework for zero-resource event extraction.

B. Data Preprocessing-Knowledge Graph Embedding

Given that the knowledge graph embedding with link prediction as the optimization goal, we evaluate the proposed model on the following for evaluation metrics: mean rank (MR), mean reciprocal rank (MRR), hit@10, hit@3, and hit@1 as follows:

$$\mathrm{MR} = \frac{1}{Q}\sum_{i=1}^{|Q|}\mathrm{rank}_i \tag{12}$$

$$\mathrm{MRR} = \frac{1}{Q}\sum_{i=1}^{|Q|}\frac{1}{\mathrm{rank}_i} \tag{13}$$

TABLE II
DETAILS OF THE DATASET SAEE

Event Type	Core arguments	Optional arguments
Enrol, Assume a position, Graduate	Actor,Time,Action,Orgization	-
Publish academic achievements	Actor,Time,Publish,Title,Journal/Conference	-
Obtain award, honor, degree	Actor,Time,Obtain,Award	-
Participate in academic activity	Actor,Time,Visit,Place	Conference
Organize academic activity	Actor,Time,Conduct,Place,Activity	-
Make academic presentation	Time,Make,Place,Title	Actor
Issue notice, list	Organization,Time,Publish,Notice	-
Pass examination, dissertation defense	Actor,Time,Pass,Exam	Place
Establishment of Organization	Organization,Time,Establish	Place
Paper published	Article,Time,Publish,Journal/Conference	-

TABLE III
DATASET COMPARISON

Corpus	Scholar academic feature vectors	Text source	Files of numbers
ACE2005 Chinese	-	general News and Weibo	633
SAEE	60	academic we-media blog	800

TABLE IV
EXPERIMENTAL RESULTS ON THE WN18RR FB15K-237 AND YAGO3-10-DR DATASETS

	WN18RR			FB15k-237			YAGO3-10-DR		
	MRR	MR	H@10	MRR	MR	H@10	MRR	MR	H@10
DisMult	.264	3798.1	.462	.151	566.3	.303	.192	5553.2	-
ConvE	.261	5007.3	.479	.154	481.7	.286	.204	4453.3	-
RotatE	.306	3374	.53	.169	333.4	.317	.214	3084.2	-
CoNE-TransE	.227	3542	.502	.306	211	.485	-	-	-
KGE (Proposed Method)	.310	3817	.531	.166	499.3	.418	.256	4688.1	-

$$\mathrm{HR@K} = \frac{\mathrm{NumberOfHits@K}}{\mathrm{GT}} \tag{14}$$

where rank_i is the order of the standard answers in the results given by the evaluated system, GT is the collection of all tests, and NumberOfHits@K represents the sum of the number of test sets in each $\mathrm{top} - K$ list.

Table IV lists the results on three public datasets in terms of three evaluation metrics. All baseline values are taken from the scores in this article [39]. On WN18RR, KGE is the best model in terms of MRR and H@10. MR is also close to the best scores. On FB15K-237, KGE performs better than most baseline. H@10 is close to the best scores. On YAGO3-10-DR, KGE is the best model in terms of MRR and other indicators are not far from the optimal indicators. From the comparison of the experiments results, we could see that the results of KGE in the hit scores are excellent on the three datasets. The higher the hit value, the higher the link predicted ranking of the entity. This is also in line with our original intention of expressing the feature of different entities through knowledge graph embedding.

C. Academic Event Extraction

For event extraction tasks, we choose event detection, event classification, event argument identification, and event

TABLE V
EXPERIMENTAL RESULTS ON THE SAEE DATASET

Model	Event identification			Event classification			Argument identification			Argument classification		
	P	R	F	P	R	F	P	R	F	P	R	F
JointBeam	-	-	-	57.31	50.66	53.78	51.79	46.87	49.21	43.53	40.64	42.04
DMCNN	69.34	58.47	63.44	63.67	55.71	59.42	53.12	47.83	50.34	52.26	40.86	45.86
JEE	49.46	57.43	53.15	47.37	50.13	48.71	61.40	50.26	55.27	55.79	46.70	50.84
RBPB	-	-	-	63.81	57.30	60.38	52.66	48.43	50.46	46.44	37.28	41.36
JMEE	71.26	60.37	65.36	65.88	58.73	62.10	67.49	60.58	63.85	58.36	52.76	55.40
Zero-Short Transfer	76.45	65.93	70.80	71.26	36.67	48.42	51.57	49.66	50.60	48.38	46.10	47.21
Joint Transition	78.49	69.49	73.72	71.52	53.27	61.06	77.31	63.48	69.72	69.73	53.44	60.51
Cross-lingual Transfer	-	-	-	71.88	50.27	61.19	-	-	-	65.18	52.39	58.09
BiLSTM-CRF attention	78.67	66.19	71.89	70.77	37.5	49.02	78.08	70.81	74.27	72.42	74.40	73.40
MIP BiLSTM-CRF attention(proposed method)	81.22	73.23	77.05	72.14	41.07	52.08	79.70	71.47	75.36	73.05	75.18	74.10

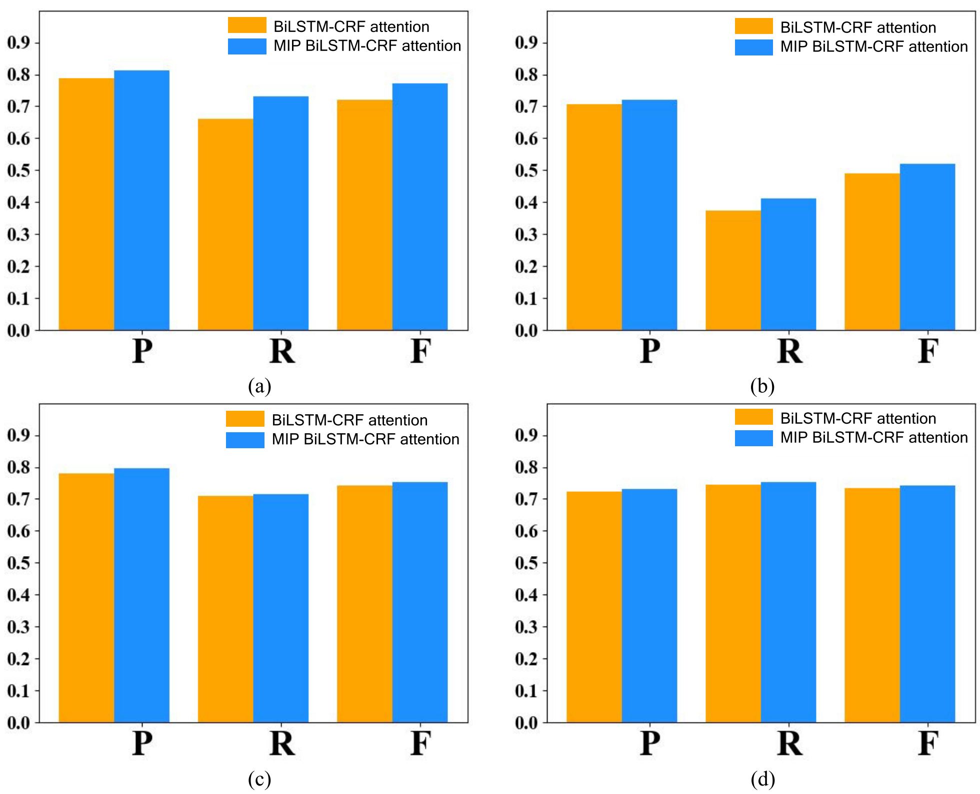


Fig. 5. Histogram of P, R, and F values about BiLSTM-CRF attention and MIP BiLSTM-CRF attention for the four tasks of event extraction. (a) Event identification. (b) Event classification. (c) Argument identification. (d) Argument classification.

argument role classification as the measurement indicators. Recall rate, precision rate, and F score could measure the quality of the model in terms of accuracy, coverage, and comprehensiveness. The calculation formulas are shown in the following equations:

$$R = \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} c_i} \tag{15}$$

$$P = \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i} \tag{16}$$

$$F = \frac{2 \times P \times R}{P + R}. \tag{17}$$

In the event extraction experiment, we use the dataset from SCHOLAT to complete. The proportion of test set, validation set, and training set is allocated according to 80%, 10%, and 10%, respectively. Since the event extraction experiment

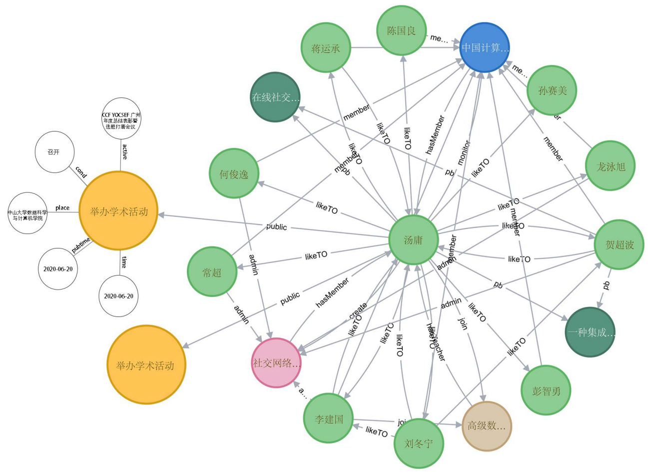


Fig. 6. Example of scholar entity in SKG.

uses the dataset published in this article, we complete the comparative experiment by searching and modifying the code published in the previous paper. The experimental results are reported in Table V. Table V lists the four aspects of event identification, event classification, argument identification, and argument classification based on the MIP BiLSTM-CRF attention method proposed in this article and the BiLSTM-CRF attention method without data preprocessing.

In the field of identification, our method has achieved significant improvements in P, R, and F. In the field of event classification, our method has improved the classification accuracy rate and decreased the recall rate. It also has a good performance in argument identification and argument classification. This indicates that our method is more effective in extraction event tasks from academic we-media blog texts. Analyzing the reasons, we believe that the reason for the increase in the experimental results might be that on the one hand, Bert integrates the position information and part of speech information of the text into the embedding vector. On the other hand, the BiLSTM-CRF framework could make the effective use of the past and future input features and sentence-level tagging information.

We also compare the performance in event extraction. As shown in Fig. 5, in the four tasks of event identification, event classification, argument identification, and argument classification, when we add the scholars' social embedding vector as input data, the recall rate, precision rate, and F values are all improved. Especially in event identification and event classification, the F values increased by 5.16% and 3.06%, respectively. This result proves that the feature of the scholar who edited the text plays an important role in text processing such as event extraction.

Fig. 6 shows an entity and some of the adjacent nodes in SKG. The first layer is the academic social network between the central node and other scholars, papers, teams, courses, and so on. The second layer is the academic events published by the central node in the social network.

VI. Conclusion

In this article, in order to extract events from academic we-media blogs into the academic knowledge graph, we propose a framework C-SKG to complete event extraction that integrates the academic social attributes of the scholar. The framework mainly includes two parts: data preprocessing and event extraction. After that, we add the extracted events to academic knowledge graph and establish the relationship between the scholar and the event. Using this framework, we construct an academic event knowledge graph SKG based on SCHOLAT. At the same time, we publish a dataset for Chinese event extraction. The scholar's social attributes and text are derived from the academic social network SCHOLAT. Compared with other methods, our methods have more excellent performance on the dataset published in this article.

The study in this article puts forward new ideas for knowledge graphs in the field of natural language processing such as event extraction. Building an event knowledge graph in academic social networks is of great significance to event sorting, event prediction, recommendation, search, and so on. The interesting directions for future work are: 1) exploring the relationship of each event and event attributes with the entities of SKG and 2) using this event knowledge graph to develop related applications.

References

[1] M. Mantle, S. Batsakis, and G. Antoniou, "Large scale distributed spatio-temporal reasoning using real-world knowledge graphs," *Knowl.-Based Syst.*, vol. 163, pp. 214–226, Jan. 2019.

[2] Y. Cheng, Y. Yuan, L. Chen, C. Giraud-Carrier, G. Wang, and B. Li, "Event-participant and incremental planning over event-based social networks," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 2, pp. 474–488, Feb. 2021.

[3] M. Rospocher *et al.*, "Building event-centric knowledge graphs from news," *J. Web Semantics*, vols. 37–38, pp. 132–151, Mar. 2016.

[4] C. Rudnik, T. Ehrhart, O. Ferret, D. Teyssou, R. Troncy, and X. Tannier, "Searching news articles using an event knowledge graph leveraged by Wikidata," in *Proc. Companion World Wide Web Conf.*, May 2019, pp. 1232–1239.

[5] W. Jin, M. Qu, X. Jin, and X. Ren, "Recurrent event network: Autoregressive structure inferenceover temporal knowledge graphs," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2020, pp. 6669–6683.

[6] S. Deng, H. Rangwala, and Y. Ning, "Dynamic knowledge graph based multi-event forecasting," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 1585–1595.

[7] T. S. Costa, S. Gottschalk, and E. Demidova, "Event-QA: A dataset for event-centric question answering over knowledge graphs," in *Proc. 29th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2020, pp. 3157–3164.

[8] H. L. Nguyen and J. J. Jung, "Social event decomposition for constructing knowledge graph," *Future Gener. Comput. Syst.*, vol. 100, pp. 10–18, Nov. 2019.

[9] K. Huang, M. Yang, and N. Peng, "Biomedical event extraction on graph edge-conditioned attention networks with hierarchical knowledge graphs," in *Proc. Conf. Empirical Methods Natural Lang. Process., Findings*, 2020, pp. 1277–1285.

[10] Y. Yang, Z. Wei, Q. Chen, and L. Wu, "Using external knowledge for financial event prediction based on graph neural networks," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manage.*, Nov. 2019, pp. 2161–2164.

[11] J. Wu, X. Zhu, C. Zhang, and Z. Hu, "Event-centric tourism knowledge graph—A case study of Hainan," in *Proc. Int. Conf. Knowl. Sci., Eng. Manage.*, 2020, pp. 3–15.

[12] F. Tang, J. Zhu, C. He, C. Fu, J. He, and Y. Tang, "SCHOLAT: An innovative academic information service platform," in *Proc. 27th Australas. Database Conf. (ADC), Databases Theory Appl.* (Lecture Notes in Computer Science), vol. 9877. Sydney, NSW, Australia: Springer, Sep. 2016, pp. 453–456, doi: 10.1007/978-3-319-46922-5_38.

[13] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol. (NAACL-HLT)*, 2019, pp. 4171–4186.

[14] K. Wang, Z. Shen, C. Huang, C.-H. Wu, Y. Dong, and A. Kanakia, "Microsoft academic graph: When experts are not enough," *Quant. Sci. Stud.*, vol. 1, no. 1, pp. 396–413, Feb. 2020.

[15] H. Wan, Y. Zhang, J. Zhang, and J. Tang, "AMiner: Search and mining of academic social networks," *Data Intell.*, vol. 1, no. 1, pp. 58–76, Mar. 2019.

[16] F. Zhang *et al.*, "OAG: Toward linking large-scale heterogeneous entity graphs," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2019, pp. 2585–2595.

[17] R. Wang *et al.*, "AceKG: A large-scale knowledge graph for academic data mining," in *Proc. ACM Conf. Inf. Knowl. Manage. (CIKM)*, 2018, pp. 1487–1490.

[18] J. Liu, J. Ren, W. Zheng, L. Chi, I. Lee, and F. Xia, "Web of scholars: A scholar knowledge graph," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2020, pp. 2153–2156.

[19] Q. Wang, Z. Mao, and B. Wang, "Knowledge graph embedding: A survey of approaches and applications," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 12, pp. 2724–2743, Dec. 2017.

[20] A. Bordes, N. Usunier, A. García-Durán, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," in *Proc. Adv. Neural Inf. Process. Syst.*, 2013, pp. 2787–2795.

[21] G. Ji, K. Liu, S. He, and J. Zhao, "Knowledge graph completion with adaptive sparse transfer matrix," in *Proc. 30th Conf. Artif. Intell.*, D. Schuurmans and M. P. Wellman, Eds., 2016, pp. 985–991.

[22] J. Feng, M. Huang, M. Wang, M. Zhou, Y. Hao, and X. Zhu, "Knowledge graph embedding by flexible translation," in *Proc. 15th Int. Conf.*, 2016, pp. 557–560.

[23] B. Yang, W. Yih, X. He, J. Gao, and L. Deng, "Embedding entities and relations for learning and inference in knowledge bases," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–12.

[24] T. Dettmers, P. Minervini, P. Stenetorp, and S. Riedel, "Convolutional 2D knowledge graph embeddings," in *Proc. 32nd AAAI Conf. Artif. Intell. (AAAI)*, 2018, pp. 1811–1818.

[25] S. Vashishth, S. Sanyal, V. Nitin, N. Agrawal, and P. P. Talukdar, "InteractE: Improving convolution-based knowledge graph embeddings by increasing feature interactions," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 3009–3016.

[26] R. Huang and E. Riloff, "Bootstrapped training of event extraction classifiers," in *Proc. 13th Conf. Eur. Chapter Assoc. Comput. Linguistics (EACL)*, 2012, pp. 286–295.

[27] K. Cao, X. Li, W. Ma, and R. Grishman, "Including new patterns to improve event extraction systems," in *Proc. 31st Int. Florida Artif. Intell. Res. Soc. Conf. (FLAIRS)*, 2018, pp. 487–492.

[28] X. Wang *et al.*, "HMEAE: hierarchical modular event argument extraction," in *Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP)*, Hong Kong, 2019, pp. 5776–5782.

[29] Y. Chen, T. Chen, S. Ebner, A. S. White, and B. Van Durme, "Reading the manual: Event extraction as definition comprehension," in *Proc. 4th Workshop Structured Predict. (NLP)*, 2020, pp. 74–83.

[30] D. Li, L. Huang, H. Ji, and J. Han, "Biomedical event extraction based on knowledge-driven tree-LSTM," in *Proc. 2019 Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol. (NAACL-HLT)*, 2019, pp. 1421–1430.

[31] H. Lin, Y. Lu, X. Han, and L. Sun, "Nugget proposal networks for Chinese event detection," in *Proc. 56th Annu. Meeting Assoc. Comput. Linguistics*, 2018, pp. 1565–1574.

[32] S. Zheng, W. Cao, W. Xu, and J. Bian, "Doc2EDAG: An end-to-end document-level framework for Chinese financial event extraction," in *Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP)*, 2019, pp. 337–346.

[33] N. Ding, Z. Li, Z. Liu, H. Zheng, and Z. Lin, "Event detection with trigger-aware lattice neural network," in *Proc. Conf. Empirical Methods Natural Lang. Process. 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP)*, Hong Kong, 2019, pp. 347–356.

[34] Z. Wang, L. Sun, X. Li, and L. Wang, "Event extraction via DMCNN in open domain public sentiment information," in *Proc. 6th Int. Conf. Pioneering Comput. Scientists, Eng. Educators (ICPCSEE), Data Science* (Communications in Computer and Information Science) vol. 1258, P. Qin, H. Wang, G. Sun, and Z. Lu, Eds. Taiyuan, China: Springer, Sep. 2020, pp. 90–100, doi: 10.1007/978-981-15-7984-4_7.

[35] M. T. Chau, D. Esteves, and J. Lehmann, "A neural-based model to predict the future natural gas market price through open-domain event extraction," in *Proc. 1st Int. Workshop Cross-Lingual Event-Centric Open Analytics Co-Located 17th Extended Semantic Web Conf. (ESWC)*, vol. 2611, 2020, pp. 17–31.

[36] Y. Dai, S. Wang, N. N. Xiong, and W. Guo, "A survey on knowledge graph embedding: Approaches, applications and benchmarks," *Electronics*, vol. 9, no. 5, p. 750, May 2020.

[37] A. Rossi, D. Barbosa, D. Firmani, A. Matinata, and P. Merialdo, "Knowledge graph embedding for link prediction: A comparative analysis," *ACM Trans. Knowl. Discovery Data*, vol. 15, no. 2, pp. 1–49, Apr. 2021.

[38] J. Cordonnier, A. Loukas, and M. Jaggi, "On the relationship between self-attention and convolutional layers," in *Proc. 8th Int. Conf. Learn. Represent. (ICLR)*, 2020, pp. 1–18.

[39] F. Akrami, M. S. Saeef, Q. Zhang, W. Hu, and C. Li, "Realistic re-evaluation of knowledge graph completion methods: An experimental study," in *Proc. Int. Conf. Manage. Data (SIGMOD)*, D. Maier, R. Pottinger, A. Doan, W. Tan, A. Alawini, and H. Q. Ngo, Eds. Portland, OR, USA, Jun. 2020, pp. 1995–2010, doi: 10.1145/3318464.3380599.

[40] G. R. Doddington, A. Mitchell, M. A. Przybocki, L. A. Ramshaw, S. M. Strassel, and R. M. Weischedel, "The automatic content extraction (ACE) program-tasks, data, and evaluation," in *Proc. 4th Int. Conf. Lang. Resour. Eval. (LREC)*, Lisbon, Portugal, May 2004, pp. 1–4.

[41] Z. Sun, Z. Deng, J. Nie, and J. Tang, "Rotate: Knowledge graph embedding by relational rotation in complex space," in *Proc. 7th Int. Conf. Learn. Represent. (ICLR)*, 2019, pp. 1–18.

[42] K. Wang, Y. Liu, X. Xu, and Q. Z. Sheng, "Enhancing knowledge graph embedding by composite neighbors for link prediction," *Computing*, vol. 102, no. 12, pp. 2587–2606, Dec. 2020, doi: 10.1007/s00607-020-00842-5.

[43] Q. Li, H. Ji, and L. Huang, "Joint event extraction via structured prediction with global features," in *Proc. 51st Annu. Meeting Assoc. Comput. Linguistics (ACL)*, 2013, pp. 73–82.

[44] Y. Chen, L. Xu, K. Liu, D. Zeng, and J. Zhao, "Event extraction via dynamic multi-pooling convolutional neural networks," in *Proc. 53rd Annu. Meeting Assoc. Comput. Linguistics 7th Int. Joint Conf. Natural Lang. Process. Asian Fed. Natural Lang. Process. (ACL)*, 2015, pp. 167–176.

[45] T. H. Nguyen, K. Cho, and R. Grishman, "Joint event extraction via recurrent neural networks," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*, 2016, pp. 300–309.

[46] L. Sha, J. Liu, C. Lin, S. Li, B. Chang, and Z. Sui, "RBPB: Regularization-based pattern balancing method for event extraction," in *Proc. 54th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, 2016, pp. 1224–1234.

[47] X. Liu, Z. Luo, and H. Huang, "Jointly multiple events extraction via attention-based graph information aggregation," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2018, pp. 1247–1256.

[48] L. Huang, H. Ji, K. Cho, I. Dagan, S. Riedel, and C. Voss, "Zero-shot transfer learning for event extraction," in *Proc. 56th Annu. Meeting Assoc. Comput. Linguistics*, 2018, pp. 2160–2170.

[49] J. Zhang, Y. Qin, Y. Zhang, M. Liu, and D. Ji, "Extracting entities and events as a single task using a transition-based neural model," in *Proc. 28th Int. Joint Conf. Artif. Intell. (IJCAI)*, 2019, pp. 5422–5428.

[50] D. Lu *et al.*, "Cross-lingual structure transfer for zero-resource event extraction," in *Proc. 12th Lang. Resour. Eval. Conf. (LREC)*, N. Calzolari *et al.*, Eds. Marseille, France, May 2020, pp. 1976–1981. [Online]. Available: https://aclanthology.org/2020.lrec-1.243/

Chao Chang was born in Changzhi, Shanxi, China, in 1994. She is currently pursuing the Ph.D. degree in software engineering with South China Normal University, Guangzhou, China.

Her research interests mainly include knowledge graph, social network services, and big data analysis for education.

Yong Tang received the B.S. and M.Sc. degrees from Wuhan University, Wuhan, China, in 1985 and 1990, respectively, and the Ph.D. degree from the University of Science and Technology of China, Hefei, China, in 2001, all in computer science.

He is currently a Professor and the Dean of the School of Computer Science, South China Normal University (SCNU), Guangzhou, China. He serves as the Director of the Services Computing Engineering Research Center of Guangdong Province. He was the Vice Dean of the School of Information of Science and Technology, Sun Yat-sen University, Guangzhou, before he joined SCNU in 2009. He has published more than 200 articles and books. As a supervisor, he has had more than 40 Ph.D. students and Post-Doctoral Researcher since 2003 and more than 100 master's students since 1996.

Dr. Tang is a Distinguished Member and the Director of the Technical Committee on Collaborative Computing of China Computer Federation (CCF). He has served as the general or program committee co-chair for more than ten conferences.

Yongxu Long is currently pursuing the master's degree in electronic information with South China Normal University, Guangzhou, China.

His current research interests are in the areas of knowledge graph, machine learning, and social networks.

Kun Hu is currently pursuing the master's degree in software engineering with South China Normal University, Guangzhou, China.

His research interests mainly include knowledge graph and software development.

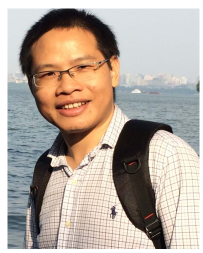
Ying Li received the M.Sc. degree from the School of Computer Science and Engineering, South China University of Technology, Guangzhou, China, in 2004. She is currently pursuing the Ph.D. degree with South China Normal University, Guangzhou.

Her research interests are community detection, machine learning, and complex network analysis.

Jianguo Li received the B.S. degree in information and computing sciences from the China University of Petroleum (East China), Qingdao, Shandong Province, China, in 2005 and the Ph.D. degree in computer software and theory from Sun Yat-sen University, Guangzhou, China, in 2010.

He is currently an Associate Professor with the School of Computer Science, South China Normal University (SCNU), Guangzhou. He has published more than 20 articles. As a supervisor, he has had more than ten master's students.

Dr. Li is an Executive Member of the Technical Committee on Collaborative Computing of China Computer Federation (CCF).



Chang-Dong Wang received the Ph.D. degree in computer science from Sun Yat-sen University, Guangzhou, China, in 2013.

He joined Sun Yat-sen University in 2013, where he is currently an Associate Professor with the School of Computer Science and Engineering. He has published over 80 scientific papers in international journals and conferences, such as IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, IEEE TRANSACTIONS ON CYBERNETICS, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, KDD, AAAI, IJCAI, CVPR, ICDM, CIKM, and SDM. His current research interests include machine learning and data mining.

Dr. Wang's ICDM 2010 paper won the Honorable Mention for Best Research Paper Awards. He received the 2012 Microsoft Research Fellowship Nomination Award. He was awarded the 2015 Chinese Association for Artificial Intelligence (CAAI) Outstanding Dissertation. He is an Associate Editor of *Journal of Artificial Intelligence Research* (JAIR).