Event Factuality Detection in Discourse

Rongtao Huang[1], Bowei Zou[1,2](✉), Hongling Wang[1], Peifeng Li[1] and Guodong Zhou[1]

[1] Natural Language Processing Lab, Soochow University, Suzhou, China
[2] Institute for Infocomm Research, Singapore
rthuang.suda@gmail.com,{zoubowei,hlwang,pfli,gdzhou}@suda.edu.cn

Abstract. Event factuality indicates whether an event occurs or the degree of certainty described by authors in context. Correctly identifying event factuality in texts can contribute to a deep understanding of natural language. In addition, event factuality detection is of great significance to many natural language processing applications, such as opinion detection, emotional reasoning, and public opinion analysis. Existing studies mainly focus on identifying event factuality by the features in the current sentence (e.g. negation or modality). However, there might be many different descriptions of factuality in a document, corresponding to the same event. It leads to conflict when identifying event factuality only on sentence level. To address such issues, we come up with a document-level approach on event factuality detection, which employs Bi-directional Long Short-Term Memory (BiLSTM) neural networks to learn contextual information of the event in sentences. Moreover, we utilize a double-layer attention mechanism to capture the latent correlation features among event sequences in the discourse, and identify event factuality according to the whole document. The experimental results on both English and Chinese event factuality detection datasets demonstrate the effectiveness of our approach. The performances of the proposed system achieved 86.67% and 86.97% of F1 scores, yielding improvements of 3.24% and 4.78% over the state-of-the-art on English and Chinese datasets, respectively.

Keywords: Event Factuality, Discourse Information, BiLSTM, Attention Mechanism.

1 Introduction

Text-oriented event factuality measures whether an event has occurred or the degree of certainty described by authors in context. Event factuality detection in texts can contribute to a deep understanding of natural language. In addition, it is of great significance for many natural language processing applications, such as question answering [1], opinion detection [2], emotion analysis [3] and rumor monitoring [4].

Event factuality is generally measured and represented by its polarity and modality. Polarity indicates whether an event has occurred in context, while modality conveys the degree of certainty. The intersection of the two dimensions produces four types of event factuality, that is, *CerTain Positive* (CT+), *CerTain Negative* (CT-), *PoSsible*

Positive (PS+), *PoSsible Negative* (PS-). Besides, if an event's factuality cannot be identified, we usually label it as *Underspecified* (Un).

新浪体育讯 这个夏天AC米兰的股权交易成为一场大戏，此前中央电视台财经频道的《环球财经连线》节目援引路透社的报道称：百度集团将以4.37亿美元**收购**（**PS+**）AC米兰。

Sina Sports News This summer AC Milan's equity transaction became a big show. Previously, the CCTV's "Global Finance Connection" program quoted Reuters as reporting that Baidu Group will ***acquire (PS+)*** *AC Milan for $437 million.*

在15日的《环球财经连线》节目中，央视称："目前，百度总裁李彦宏与意大利AC米兰的谈判已经有了进展，预计将以4亿3700万美元**收购**（**PS+**）AC米兰。"

In the "Global Finance Connection" program on the 15th, CCTV said: "At present, the negotiation between Baidu President Li Yanhong and Italy's AC Milan has progressed, and it is expected to ***acquire (PS+)*** *AC Milan for $437 million."*

这消息一出，就引来一片质疑之声，因为AC米兰80%股份估值5亿欧元，如果接盘，还需要考虑2亿欧元的债务，那么4.37亿美元，约3.93亿欧元**收购**（**PS-**）AC米兰的消息看起来有些禁不起推敲。

When this news came out, it led to a voice of doubt because AC Milan's 80% stake was valued at 500 million euros. If it took over, Baidu still needs to consider 200 million euros of debt. Thus the news of ***acquiring (PS-)*** *AC Milan for 437 million dollars, about 393 million euros does not seem to work.*

据新浪消息，19日早间，百度方面否定其参与**收购**（**CT-**）意甲俱乐部AC米兰。此前央视报道称，百度已完成了这一4.37亿美元的**收购**（**CT+**）计划。

According to Sina News, on the morning of the 19th, Baidu denied its participation in the ***acquisition (CT-)*** *of Serie A club AC Milan. Previously, CCTV reported that Baidu had completed the $437 million* ***acquisition (CT+)*** *plan.*

Fig. 1. An example of event factuality (Bold: Event).

In a document, there might be different descriptions of factuality in different sentences, corresponding to the same event. As shown in Fig. 1, the acquisition event appears five times in a document, ignoring the contextual information and judging event factuality only based on individual sentence in which each event is located. Among them, the first and second times are CCTV's speculation on the occurrence of this event, with factuality PS+. At the third time, the author questioned the occurrence of the acquisition event by "看起来有些禁不起推敲(*does not seem to work*)", thus its factuality is PS-. Obviously, at the fourth time, Baidu, the participant in the acquisition event, explicitly denied the occurrence of the event via a negative cue "否定(*denied*)", so its factuality is CT-. Finally, for the last mention, it is judged from the sentence perspective with factuality CT+. However, for the same event, whether it has occurred or not can only be one situation (positive/negative). Furthermore, in the same document, the degree of certainty of the event ultimately comes down to one attitude (certainty/possible). Based on the analysis in Fig. 1, as a direct participant in the acquisition event, Baidu clearly denied the occurrence of the event (the fourth mention), thus CT- is inferred as the document-level factuality of the event.

Existing studies on event factuality detection usually focus on the sentence level. Cao et al. proposed a three-dimensional representation system, expressing event factuality as a triple of <polarity, level, tense> [5]. Qian et al. extracted event triggers, event sources, negative and speculative cues from raw texts [6]. However, compared to document-level event factuality in Fig. 1, sentence-level event factuality easily leads to conflicts between different mentions of the same event, which makes it difficult to apply to NLP tasks such as information extraction and knowledge base construction. In addition, according to statistics on the English and Chinese event factuality datasets, 25.4% (English) and 37.8% (Chinese) of instances are inconsistent between sentence-level factuality and document-level factuality for the same event mention.

To address the above issue, we propose a document-level event factuality detection approach. Specifically, we employs BiLSTM networks to learn the contextual information of the event in sentences. Such BiLSTM feature encoder can effectively model the forward and backward information around the event. Then, we come up with a double-layer attention mechanism to capture the latent correlation features among event sequences in discourse. In particular, first, the intra-sequence attention mechanism can capture the dependence between cues and event in the sentence. Second, the inter-sequence attention mechanism can extract the document-level feature representation of the event from the event sequence. Finally, the probability of the event factuality is decoded by a softmax layer.

The experimental results on both English and Chinese event factuality detection datasets [7] demonstrate the effectiveness of our approach. The performances of the proposed system achieve 86.67% and 86.97% of F1 scores, yielding improvements of 3.24% and 4.78% over the state-of-the-art on English and Chinese datasets, respectively. In addition, the related experiments also verify the effectiveness of event triggers, negative and speculative cues on document-level event identification.

2 Related Work

Early studies on event factuality detection concentrated on the sentence level. Minard et al. released the MEANTIME corpus [8] and analyzed that the event factuality is characterized by certainty, tense and polarity. According to their theory, certainty includes three subcategories of "certainty", "uncertainty", and "unspecified"; Tense distinguishes among "past", "future", and "unspecified"; And polarity is divided into "positive", "negative", and "unspecified". Besides, Minard proposed an event factuality detection system, FactPro [9,10]. Saurí et al. released the FactBank corpus [11] and divided factuality values into seven categories according to the modality and polarity of the event, i.e. Fact (CT+), Counterfact (CT-), Probable (PR+), Not probable (PR-), Possible (PS+), Not possible (PS-) and underspecified (Uu). Moreover, Saurí proposed the De Faco system [12], which traverses the dependency syntax tree of the event from top to bottom, and calculates the factuality of the event layer by layer. Recently, neural networks are effectively applied to various NLP tasks. Qian et al. [6] extracted event factuality information from raw text and proposed a generative adversarial network with auxiliary classification for event factuality detection.

In Chinese, Cao Yuan [13] annotated the Chinese event factuality based on the ACE 2005 corpus and proposed a 3D representation system, regarding the event factuality as a <polarity, level, tense> triplet. On the basis, He et al. [14] proposed a Convolutional Neural Network (CNN) based Chinese event factuality detection model.

However, all of the above studies focus on identifying event factuality by relevant features (e.g. negative and speculative cues) in sentence level. Qian annotated an event factuality data set for Chinese and English news texts in his PhD thesis [7], which marked the event factuality on both document level and sentence level, and proposed a document-level event factuality detection method based on adversarial networks.

3 Document-level Event Factuality Detection

In this paper, we propose a document-level event factuality detection approach that comprehensively considers effective information related to the target event in a document. First, a BiLSTM neural network is employed to learn contextual information of the target event. Then, we utilize a double-layer attention mechanism to capture latent correlation features among the event sequence in discourse. Fig. 2 illustrates the framework for our event factuality detection approach.

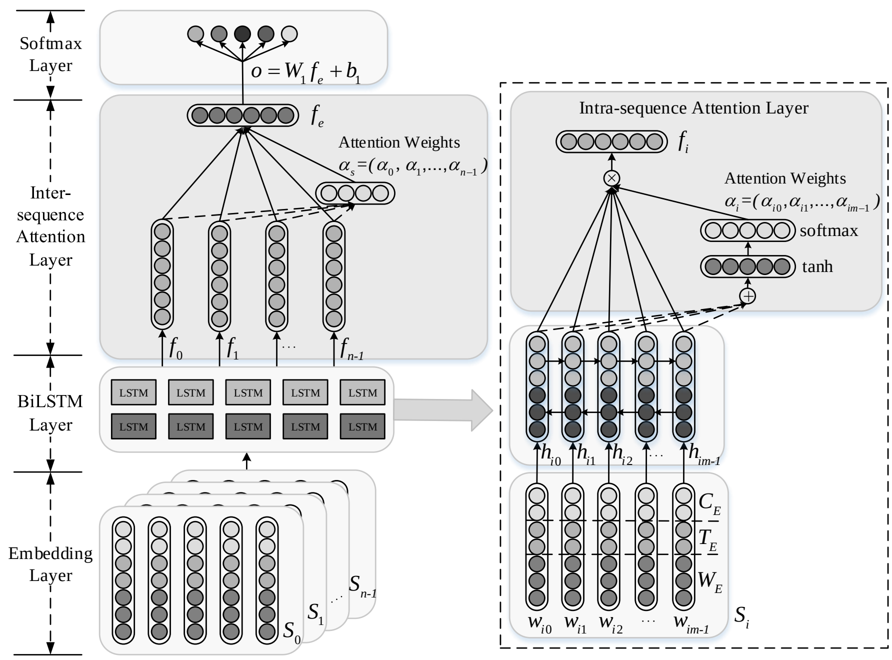

Fig. 2. Framework of event factuality detection model in discourse.

3.1 Embedding Layer

First, we encode the sentence sequence that contains the target event and corresponding features in discourse. Specifically, given the target event E, assume that the sentence sequence containing E is $(S_0, S_1, \dots, S_{n-1})$, where $S_i = (w_{i0}, w_{i1}, \dots, w_{im-1})$, m is the length of S_i. We transform each word w_{ij} into a real-valued vector with dimension d_w by using a word embedding matrix $W_E \epsilon \mathbb{R}^{d_w \times |V|}$, where V is the input vocabulary.

Event Trigger: We transform each trigger tag into a vector with the dimension d_t by a matrix $T_E \epsilon \mathbb{R}^{d_t \times |V_t|}$, where V_t is the set of trigger tags, $V_t = \{0,1\}$, 1 denotes an event trigger, while 0 indicates a non-trigger.

Negative and speculative cue: Similarly, we transform each cue tag into a vector with the dimension d_c by a matrix $C_E \epsilon \mathbb{R}^{d_c \times |V_c|}$, where V_c is the set of cue tags, $V_c = \{0,1,2\}$, 1 denotes a negative cue, and 2 indicates a speculative cue, while 0 represents a non-cue.

Finally, we represent the sentence sequence as a matrix $\mathrm{X} \in \mathbb{R}^{d_0 \times m}$, where $d_0 = d_w + d_t + d_c$, m is the length of the sequence.

3.2 BiLSTM Layer

To capture the contextual information of the target event in a sentence, we employ BiLSTM [15] networks to learn the forward representation $\vec{H}$ and the backward representation $\overleftarrow{H}$ of the sentence. Then, the characteristic representation of the target event $H \in \mathbb{R}^{m \times n_h}$ is obtained by splicing $\vec{H}$ and $\overleftarrow{H}$, where $n_h = 2 \times {n_h}^*$, and ${n_h}^*$ indicates the number of the hidden layer units in the BiLSTM.

$$H = \vec{H} \oplus \overleftarrow{H} \tag{1}$$

3.3 Intra-sequence Attention Layer

We employ an intra-sequence attention mechanism [16] to learn the weight distribution of each element in the sentence, and combine the information according to the weight distribution to acquire the characteristic representation of the event $f \in \mathbb{R}^{n_h}$ in the sequence:

$$H_m = \tanh(H) \tag{2}$$

$$\alpha = \mathrm{softmax}\left(v \cdot {H_m}^T\right) \tag{3}$$

$$f = \tanh(\alpha \cdot H) \tag{4}$$

where tanh is the hyperbolic tangent function, "·" denotes the point multiplication operation, and $v \in \mathbb{R}^{n_h}$ are model parameters.

3.4 Inter-sequence Attention Layer

Given a target event, suppose that there are n sentences including the target event in the document, the sentence sequence can be represented as $X = (X_0, X_1, \dots, X_{n-1})$, and the corresponding features are $F_s = (f_0, f_1, \dots, f_{n-1})$, where $f_i = f$. To acquire the importance of different sentences on the document-level event factuality, we similarly utilize the attention mechanism to assign different weights to different sentences, and combine the sentence-level features according to the weight distribution to acquire the document-level characteristic representation of the target event $f_e \in \mathbb{R}^{n_h}$:

$$H_{ms} = \tanh(F_s) \quad (5)$$

$$\alpha_s = \mathrm{softmax}\left(v_s \cdot {H_{ms}}^T\right) \quad (6)$$

$$f_e = \tanh(\alpha_s \cdot F_s) \quad (7)$$

3.5 Softmax Layer

Event factuality detection is essentially a classification task, so we utilize a softmax layer as the classifier. The input of softmax layer is the document-level event characteristic representation f_e, and the output is the probability of the factuality values, as follows:

$$o = \mathrm{softmax}(W_1 f_e + b_1) \quad (8)$$

where $W_1 \in \mathbb{R}^{c \times n_h}$, $b_1 \epsilon \mathbb{R}^c$ are model parameters, and c is the number of event factuality values. We also employ the cross-entropy cost to measure the error between the predicted value and the true value.

4 Experimentation

This section introduces experimental datasets, evaluation metrics, experimental tools and parameter settings. Then we show experimental results and demonstrate the effectiveness of the proposed approach and features.

4.1 Experimental Settings

In this paper, we adopt the English and Chinese event factuality datasets [7], which annotated the event factuality on both document-level and sentence-level. The number of English and Chinese documents are 1,730 and 4,650, respectively, which is from China Daily, Sina Bilingual News, and Sina News. Table 1 lists the distribution of event factuality categories. From lines 1-2, we can see that the *certain positive* (CT+) category includes the largest number of instances, accounting for 66.5% (English) and 51.7% (Chinese), while *possible negative* (PS-) and *underspecified* (Un) are only about 1%. Therefore, we mainly evaluate and compare the performances of the system in the CT+, CT- and PS+.

In addition, to find the difference of event factuality between sentence-level and document-level, we statistic the number of documents that meet the following conditions: for the same event, there are n sentences in the document whose factuality is different from the document-level (lines 3-7 in Table 1). We can see that 1) in 25.4% (English) and 37.8% (Chinese) of documents, the annotations of the sentence-level factuality of the same event are inconsistent with the document-level, which indicates that identifying event factuality only on sentence-level may lead to conflict; and 2) in such documents, there are more CT- and PS+ categories (document-level) and fewer CT+ categories.

Table 1. Statistics of dataset (#document).

Items	Chinese						English					
	CT+	CT-	PS+	PS-	Un	Total	CT+	CT-	PS+	PS-	Un	Total
Discourse	2,403	1,342	848	36	21	4,650	1,150	279	274	12	15	1,730
Sentence	11,487	3,924	2,879	123	593	19,006	4,401	662	574	37	81	5,575
n=0	2,066	487	319	9	9	2,890	1,026	164	93	2	5	1,290
n=1	231	390	269	10	5	905	108	56	91	5	4	264
n=2	68	217	159	9	4	457	12	28	54	2	2	100
n=3	17	126	54	2	1	200	1	15	22	1	1	40
n≥4	21	122	47	6	2	198	8	17	15	2	0	36

We use a fixed 80%/10%/10% split for training, developing, and testing, respectively. For measurement, traditional Precision, Recall, and F1-score are used to evaluate the performance in event factuality detection. In addition, the Macro-Averaging and Micro-Averaging is also adopted to evaluate the average performance over three factuality categories from different aspects.

The Chinese negative and speculative cues adopt in this paper are annotated in the CNeSp[1] corpus [17], and the English are from the BioScope[2] corpus [18]. We employ ELMo[3] as the pre-trained word embeddings with the dimension 1,024. In our experiment, the dependency syntax paths of the Chinese are generated by the Chinese language processing toolkit[4], and the part-of-speech and dependency syntax paths of English are generated by the Stanford CoreNLP[5]. Besides, we set the hidden units in LSTM $n_h{}^* = 100$ and the dimension of the event triggers, negative and speculative cues as 100 and 200, respectively. Other parameters are initialized randomly, and all the models are optimized using the stochastic gradient descent (SGD) with momentum.

To verify the effectiveness of our approach, we compare several baselines on event factuality detection, which are briefly introduced as follows.

BiLSTM: The BiLSTM model with word embeddings.

BiLSTM+Att: The attention-based BiLSTM model with word embeddings.

BiLSTM+Att_E: The attention-based BiLSTM model with word embeddings and event trigger embeddings.

BiLSTM+Att_C: The attention-based BiLSTM model with word embeddings and negative and speculative cues embeddings.

BiLSTM+Att_E_C: The attention-based BiLSTM model with word embeddings, event trigger embeddings, negative and speculative cues embeddings.

Att+Adv: The document-level approach [7] based on adversarial networks with the dependency syntax path of the cues in adjacent sentences to the target event.

1 `http://nlp.suda.edu.cn/corpus/CNeSp/`
2 `http://www.inf.u-szeged.hu/rgai/bioscope`
3 `https://github.com/HIT-SCIR/ELMoForManyLangs`
4 `http://hlt-la.suda.edu.cn`
5 `https://stanfordnlp.github.io/CoreNLP/index.html`

BiLSTM+Att+J: The joint learning model, which adds an identical network structure (sentence-level channel) to the framework shown in Fig. 2 (document-level channel). In the sentence-level channel, only the current sentence including the target event is considered, and features are the same as the document-level channel.

4.2 Results and Analysis

Effect of various models. Tables 2 and 3 show the performances of systems on event factuality detection. We can see that, our approach (BiLSTM+Att_E_C) achieved 86.67% and 86.97% of F1 scores (Micro-Average), yielding improvements of 3.24% and 4.78% over the state-of-the-art on English and Chinese datasets, respectively, which indicates that comparing with the Att-Adv, our approach is simpler and more effective. It does not need to use syntactic structure information to avoid introducing feature-level noise and enhance generalization.

In addition, the comparative tests show that: 1) On all models, the CT+ category has the highest performance, followed by CT-, and PS+ has the lowest. The main reason is that there are more authentic, less rigorous or false reports in news texts; 2) the BiLSTM+Att model outperforms the BiLSTM model, with 2.86% gain in Macro-Average, which indicates that attention mechanism can capture the latent correlations among the event sequence and demonstrates the effectiveness of attention mechanism for this task; 3) the performances of the joint learning model are slightly lower than the best. The reason may be that sentences in the document have different factuality descriptions of the same event, and there is no inevitable connection with the document-level event factuality.

Table 2. Performances on English event factuality detection (P%/R%/F1%).

Models	CT+	CT-	PS+	Macro-Average	Micro-Average
BiLSTM	76.51/93.18/84.00	64.93/58.06/60.49	58.17/22.23/31.78	66.54/57.82/61.87	73.12/75.30/74.19
BiLSTM+Att	81.11/93.64/86.91	71.13/59.68/64.81	74.48/53.71/62.40	75.57/69.01/72.13	78.84/80.95/79.88
BiLSTM+Att_E_C	90.45/90.08/90.26	**78.68/88.21/83.14**	**78.88/77.41/78.06**	**82.67/85.23/83.93**	**85.77/87.59/86.67**
Att+Adv (Qian)	87.28/91.18/83.25	80.57/76.26/77.82	66.81/60.98/62.61	78.22/76.14/76.49	82.77/83.75/83.25
BiLSTM+Att +J	**86.17/95.91/90.74**	80.99/77.42/78.82	87.74/66.66/75.75	84.97/79.99/82.38	85.26/87.80/86.51

Table 3. Performances on Chinese event factuality detection (P%/R%/F1%).

Models	CT+	CT-	PS+	Macro-Average	Micro-Average
BiLSTM	76.58/88.03/81.91	80.73/72.43/76.33	71.57/57.22/63.59	76.29/72.56/74.38	76.89/77.39/77.14
BiLSTM+Att	80.07/89.10/84.33	85.26/76.47/80.59	72.96/63.89/68.10	79.43/76.49/77.92	80.17/80.43/80.30
BiLSTM+Att_E_C	**89.00/91.71/90.28**	84.64/86.34/85.42	82.09/75.18/78.35	85.24/84.41/84.81	**86.45/87.49/86.97**
Att+Adv (Qian)	83.89/89.33/86.49	80.96/79.79/80.30	77.08/67.12/71.44	80.64/78.75/79.41	81.94/82.45/82.19
BiLSTM+Att +J	87.12/92.52/89.74	**84.76/87.50/86.09**	**86.26/72.23/78.52**	**86.05/84.08/85.05**	86.22/87.07/86.64

Effect of different features. We discover that the event triggers in English datasets are mostly verbs and the corresponding part-of-speech can represent the tense to a certain degree. Therefore, we added the part-of-speech of the trigger as the tense feature to the model.

BiLSTM+Att_E_C_P: The attention-based BiLSTM model with word embeddings, event trigger embeddings, cue embeddings, and part-of-speech of the trigger.

Dual_Path: The two-channel learning model based on the model proposed in this paper. The characteristics of channel one are the same as the BiLSTM-Att_E_C model, and the characteristics of channel two are the dependent syntax path of the negative or speculative cues in adjacent sentences to the target event.

Table 4 shows the comparison of the BiLSTM+Att models with different feature embeddings. We can see that 1) the performances are significantly improved when adding the characteristics of event triggers, negative and speculative cues. It indicates that such features provide the obvious help for event factuality identification; 2) when adding the part-of-speech features, the performances are slightly reduced in English datasets, the reason may be that the part-of-speech can not effectively represent the tense of events; 3) when adding the Dependent syntax path, the performances are reduced, which may be that there is no semantic connection between the cues in adjacent sentences and the target event.

Table 4. Effect of features on event factuality detection (P%/R%/F1%).

Features	English		Chinese	
	Macro-Average	Micro-Average	Macro-Average	Micro-Average
BiLSTM+Att	75.57/69.01/72.13	78.84/80.95/79.88	79.43/76.49/77.92	80.17/80.43/80.30
BiLSTM+Att_E	76.07/75.12/75.59	81.31/83.73/82.50	79.87/76.64/78.21	80.44/81.09/80.76
BiLSTM+Att _C	81.89/79.92/80.88	83.59/85.91/84.73	82.46/82.94/82.70	83.60/84.20/83.90
BiLSTM+Att _E_C	**82.67/85.23/83.93**	**85.77/87.59/86.67**	**85.24/84.41/84.81**	**86.45/87.49/86.97**
BiLSTM+Att _E_C_P	84.10/78.87/81.39	84.75/87.10/85.91	N/A	N/A
Dual_Path	81.94/74.46/77.97	82.37/84.82/83.57	81.28/79.38/80.30	81.93/82.28/82.11

5 Conclusion

In this paper, we propose a document-level approach on event factuality detection, which employs BiLSTM neural networks to learn contextual information of event in sentences. Moreover, we utilize a double-layer attention mechanism to capture the latent correlation features among the event sequence in the discourse and identify event factuality according to the whole document. Experiments on both English and Chinese event factuality detection datasets demonstrate the effectiveness of our approach. In the future, we will explore how to better extract and represent the tense and the source of events. On the other hand, self-attention mechanisms can effectively learn the internal structure information in sequence. Thus how to transfer the self-attention mechanism to the document-level event factuality detection is also needs to be explored.

Acknowledgments. This research was supported by National Natural Science Foundation of China (Grants No.61703293, No.61672368, No.61751206). The authors would like to thank the anonymous reviewers for their insightful comments and suggestions.

References

1. Saurí R, Verhagen M, Pustejovsky J. Annotating and recognizing event modality in text[C]//Proceedings of 19th International FLAIRS Conference. 2006.
2. Wiebe J, Wilson T, Cardie C. Annotating expressions of opinions and emotions in language[J]. Language resources and evaluation, 2005, 39(2-3): 165-210.
3. Klenner M, Clematide S. How factuality determines sentiment inferences[C]//Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics. 2016: 75-84.
4. Qazvinian V, Rosengren E, Radev D R, et al. Rumor has it: Identifying misinformation in microblogs[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2011: 1589-1599.
5. Cao Y, Zhu Q, Li P. 3D Representation of Chinese Event Factuality[J]. In Proceedings of the 15th Chinese Lexical Semantic Workshop. 2014. 7~13.
6. Qian Z, Li P, Zhang Y, et al. Event Factuality Identification via Generative Adversarial Networks with Auxiliary Classification[C]//IJCAI. 2018: 4293-4300.
7. Qian Z. Research on Methods of Event Factuality Identification [D]. Jiangsu: Soochow University, 2018. (in Chinese)
8. Minard A L, Speranza M, Urizar R, et al. MEANTIME, the NewsReader multilingual event and time corpus[J]. 2016.
9. Minard A L, Speranza M, Caselli T, et al. The EVALITA 2016 Event Factuality Annotation Task (FactA)[C]//of the Final Workshop 7 December 2016, Naples. 2016: 32.
10. Minard A L, Speranza M, Sprugnoli R, et al. FacTA: Evaluation of Event Factuality and Temporal Anchoring[C]//Proceedings of the 2nd Italian Conference on Computational Linguistics. 2015: 187-192.
11. Saurí R, Pustejovsky J. FactBank: a corpus annotated with event factuality[J]. Language resources and evaluation, 2009, 43(3): 227.
12. Saurí R. A factuality profiler for eventualities in text[J]. Unveröffentlichte Dissertation, Brandeis University. Zugriff auf http://www.cs.brandeis.edu/~roser/pubs/sauriDiss, 2008, 1.
13. Cao Y, Zhu Q, Li P. The Construction of Chinese Event Factuality Corpus[J]. Journal of Chinese Information Processing, 2012, 27(6): 38-44. (in Chinese)
14. He T, Li P, Zhu Q. Approach to identify Chinese Event Factuality[J]. Journal of Chinese Information Processing, 2017, 44(5): 241-244+256. (in Chinese)
15. Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures[J]. Neural Netw, 2005, 18(5):602-610
16. Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[C]. ICLR, 2015.
17. Zou B, Zhu Q, Zhou G. Negation and speculation identification in Chinese language[C]//Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2015, 1: 656-665.
18. Vincze V, Szarvas G, Farkas R, et al. The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes[J]. BMC bioinformatics, 2008, 9(11): S9.