Rumor Detection Based on the Temporal Sentiment

Chenbo Fu[1,2], Kang Chen[1,2], Xingyu Pan[1,2], Shanqing Yu[1,2], Jun Ni[3], and Yong Min[4,5](✉)

[1] Institute of Cyberspace Security, Zhejiang University of Technology, Hangzhou 310023, China
[2] College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China
[3] Hangzhou Zhongao Technology Company Limited, Hangzhou 310000, China
[4] Computational Communication Research Center, Beijing Normal University, Zhuhai 519087, China
[5] School of Journalism and Communication, Beijing Normal University, Beijing 100875, China
myong@bnu.edu.cn

Abstract. The development of social media has changed the way of information consumption by the public, and it has also shifted the spread of rumors from offline to online. The combined multiscale nature of rumors makes it challenging to develop an effective rumor detection method. This raises the fashion of multi-modal rumor detection. However, these multi-modal methods usually focus on explicit textual and visual features, ignoring the sentiment hidden in textual content, which expresses the individuals' opinions. Thus, we propose a rumor detection model based on temporal sentiment features in this work. Specifically, we first extract the temporal sentiment feature and text vectors from the text content in normalized reply series, then combine these two vectors as the microblog representation. After that, we apply RvNN to capture the comprehensive representation of the event. Finally, we adopt the multi-layer perceptron neural network to detect rumors. The experiments on two real-world datasets, i.e., Weibo and RumourEval-2019, show that our method performs better than baseline methods. Moreover, the ablation study and the early rumor detection experiments show the effectiveness of our temporal sentiment feature. Our work supplements current rumor detection methods and highlights the important role of temporal sentiment features in rumor spreading.

Keywords: Rumor detection · Rumor spreading · Sentiment analyze · Temporal sentiment · Social network

1 Introduction

The development of social media has changed the way of information consumption by the public. However, the massive amount of information also brought many rumors [1,2]. Rumor is defined as information without official verification [3,4] and most rumors have a bad influence, for example, a lot of rumors

X. Meng et al. (Eds.): BDSC 2022, CCIS 1640, pp. 275–290, 2022.
https://doi.org/10.1007/978-981-19-7532-5_18

emerged during the 2016 US election [5] and the period of the COVID-19 pandemic [6,7]. However, the cost of rumor detection is high, especially for the manual fact-checking methods [8]. And unfortunately, research has proven that humans are only a little better than random identification when facing rumors [9]. Therefore, it is necessary to develop effective rumor detection methods. Fortunately, with the continuous development of deep learning [10–14], more and more rumor detection methods have been proposed and used to detect rumors [15–19].

Most of the transitional rumor detection methods focus on textual features [20]. Recently, as social media develops, more features are investigated and integrated as the complex representation of the rumors [17,21], e.g., visual features. However, malicious users still can easily manipulate textual content and images through technology, e.g., deep fake and text generation, and human beings hard detect these generations [22–24]. Furthermore, compared to social media's textual and visual features, the propagation features are hard to manipulate by malicious users. Thus, more and more multi-modal detection models integrate the propagation feature into their methods [25–27].

Sentiments as a representation of people's psychological state [28] have also received increasing attention from researchers [29–36]. Most existing sentiment-based rumor detection methods treat sentiment as a supplement attribute of textual features [29,35] but ignore the sentiment change when the information spreads. Recent research finds that sentiment will change with the spreading of rumors, e.g., Davoudi et al. constructed and analyzed the sentiment network by connecting posts with similar sentiment scores, and found the event sentiment changes as rumor spreading [36].

To effectively capture the sentiment feature change as the rumor spreads, we propose the temporal sentiment feature, which encodes the sentiment of the text content and the normalized reply series. By integrating the temporal sentiment feature into the microblog representation, we propose a new model to detect rumors. Extensive experiments on the two real-world datasets, i.e., Weibo and RumourEval-2019, show that our proposed method performs better than the baselines. Moreover, the ablation study and early detection experiments prove the effectiveness of our temporal sentiment feature.

The main contributions of our work are summarized as follows:

- We propose the temporal sentiment feature, which can effectively capture the sentiment feature as information spreading.

- We propose a new multi-modal rumor detection method that includes the temporal sentiment feature. The results in two real-world datasets, i.e., the Weibo and RumourEval-2019, show that our proposed method performs better than the baseline methods, and the ablation study proves the effectiveness of our temporal sentiment feature.

The rest of this paper is organized as follows: Sect. 2 introduces related work, Sect. 3 presents the details of our method, and Sect. 4 gives detailed experimental steps, results, and discussions. Finally, we summarize our work in Sect. 5.

2 Related Work

2.1 Single-Modal Rumor Detection

The existing rumor detection methods can be roughly divided into single-modal and multi-modal methods. Single-modal methods mainly focus on one type of rumor feature, i.e., statistical and embedded features from text content [16,18,37], visual features from images [38,39] and propagation features from spreading networks or social networks [15,40]. For example, Bond et al. focused on the difference of rumors in semantic-level features in text content [37], such as the uncertainty words. Kaliyar et al. proposed the FakeBERT model to detect rumors, which mainly utilized the BERT to embed the text content of the microblogs [16]. For rumor detection based on visual features, Guarnera et al. modeled the convolutional generative process and extracted a set of local features utilizing the Expectation Maximization algorithm to detect the fake images [39]. Furthermore, Zhao et al. proposed a multi-attentional deepfake detection network that consisted of multiple spatial attention heads, visual feature enhancement blocks, and attention maps to find the fake images in rumors [38]. Moreover, Ma et al. proposed the PTK model to detect rumors which evaluated the similarities of propagation tree structures between rumors [40]. Bian et al. proposed the Bi-GCN model to explore the propagation and dispersion structures of rumors with top-down and bottom-up GCN [15].

2.2 Multi-modal Rumor Detection

However, single-modal methods can not capture the multi-media nature of social media. Thus, the single-modal methods are challenged by multi-media information and lead to low accuracy. Furthermore, the development of deep fake and text generation techniques makes malicious users can fool the rumor detection models easily [22–24]. For this reason, multi-modal methods are developed to capture the multi-information, e.g., combining the textual and visual features [17,21,41]. For example, Qian et al. proposed the HMCAN model, which used the ResNet to extract features of the image and used the BERT to extract features of text content [17]. Wang et al. proposed the EANN model, which used the Text-CNN to extract the feature of text content and the VGG-19 to extract the visual feature [21]. There are also many multi-modal methods based on the textual and user features. For example, Vo et al. proposed the MAC model, which combined the user information with the text content and used the attention mechanism to capture the feature of the source post and the replies [42]. Dou et al. combined the user engagement information and the text content to detect the rumors [43]. Furthermore, multi-modal methods also exist based on the propagation structure and text content. For example, Lu et al. proposed the GCAN model, which fused the text embedding and propagation representation to detect the rumors [19].

2.3 Rumor Detection Based on Sentiment Analyze

The sentiment is a latent textual feature hidden in the text content. It has been proven that sentiment is a vital feature for rumor detection [32,33,44]. Most previous sentiment based rumor detection methods focus on the sentiment in the source microblogs, such as Mackey et al. combined sentiment and the word embedding of source microblog [30]. Yang et al. proposed the TI-CNN model to extract the explicit and latent features of microblogs, where their explicit feature of the text included the sentiment [31]. Wang et al. proposed the SD-DTS-GRU model, which focused on the fine-grained sentiment of source microblogs [32]. There are also works focusing on the sentiment of both source microblogs and replies [29,35,45]. For example, Zhang et al. combined the publisher sentiment, social sentiment, and the sentiment gap as the dual emotion features with existing rumor detection methods to detect the rumors [29]. Guo et al. proposed the EFN model, which captured both sources' and replies' textual and sentimental features for rumor detection [45]. Recent research shows that the sentiment changes as the information spreads, and has a significant difference between rumor and non-rumor [46,47]. For this reason, Davoudi et al. proposed the DSS model, which captured the features of the sentiment network and propagation tree to detect rumors [36]. Inspired by the above works, we designed the temporal sentiment feature to detect rumors in this study.

3 Method

Our work focuses on rumor detection on social networks. The architecture of our model is illustrated in Fig. 1. As shown in Fig. 1, the architecture consists of three components, i.e., Microblog Representation, Comprehensive Representation and Rumor Classifier. First, we use Microblog Representation to represent one microblog of the event, then use the RvNN to catch the feature of microblog representation along the spreading path and get the Comprehensive Representation of the event. Finally, the Rumor Classifier is used to judge whether the event is a rumor.

3.1 Problem Statement

We denote the event-based dataset as $\mathcal{C} = \{C_1, C_2, ..., C_{|\mathcal{C}|}\}$, where $|\mathcal{C}|$ represents the number of event. Each event C_i is modeled as a tree structure $C_i = \langle V_i, E_i \rangle$ where i is the index of event. And $V_i = \{t_0^i, t_1^i, t_2^i, ..., t_{k_i-1}^i\}$ is the set of nodes, where k_i is the number of microblogs of event C_i, t_j^i is the microblogs of event C_i and the j is the index of microblog which are sorted by posting order. Because the source microblog must be the first microblog of the event C_i, so the j of the source microblog is 0 and we use t_0^i represents the source microblog of event C_i. $E_i = \{e_{st}^i | s, t = 0, ..., k_i - 1\}$ is the edge set of event C_i. For example, if t_m^i is the replay of t_n^i, where the t_n^i and t_m^i represent two microblogs in event C_i, there will have a direct edge $t_n^i \rightarrow t_m^i$, i.e., e_{nm}^i. The rumor detection task can

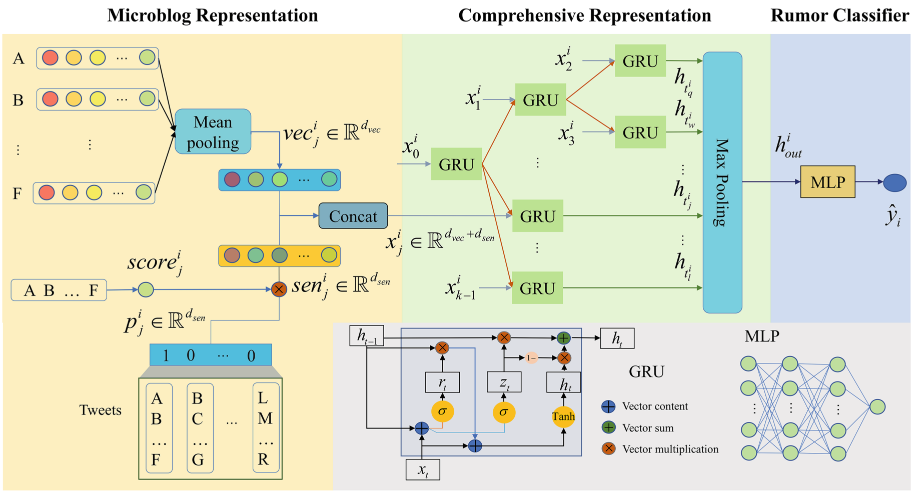


Fig. 1. The architecture of the proposed rumor detection model.

be defined as a classification problem, thus, our model outputs the $\hat{y}_i \in 0, 1$ to detect label of event C_i, where $y_i = 1$ and $y_i = 0$ denote the post is rumor and non-rumor, respectively.

3.2 Microblog Representation

Temporal Sentiment Feature. The sentiment as a representation of people's psychological state is hidden in the text content. In this work, we use the Baidu sentiment API[1] and the NLTK[2] to analyze the sentiment of the Chinese and English text, respectively. These sentiment analysis tools give a sentiment score $score^i_j$ of the text of microblog j in event i and range from -1 to 1, the score closer to 1, the more positive the text is.

To characterize the temporal features, we use a one-hot vector to encode the replies in the microblog in posting order. However, the number of responses to different social posts is heterogeneous and cannot be directly encoded, e.g., the dimension is different. Furthermore, the time interval between the two replies in the same event has a large variation, e.g., there are many replies in the early period and short time, accompanied by sporadic replies in the later period. Therefore, we need to normalize the length of the reply series, whose length is the number of posts. It should be noted that our reply series includes the source post. We use the following function to map the original one-hot vector to the normalized reply series $p^i_j \in \mathbb{R}^{d_{sen}}$:

$$p^i_j = [I(0 = \lfloor j\frac{d_{sen}}{k_i}\rfloor), I(1 = \lfloor j\frac{d_{sen}}{k_i}\rfloor), ..., I(d_{sen} - 1 = \lfloor j\frac{d_{sen}}{k_i}\rfloor)]^\top \quad (1)$$

[1] https://ai.baidu.com/tech/nlp_apply/sentiment_classify.
[2] https://www.nltk.org/.

where $\lfloor * \rfloor$ represents floor function which gives the greatest integer less than or equal to the input, $I(l = k)$ means if k is equal to l, the result will be 1, otherwise is 0. d_{sen} is the dimension of the normalized reply series, which is set as 100 in this study.

After calculating the normalized reply series, we incorporate microblog sentiment into the temporal vector. Specifically, we multiply the generated one-hot vector with the sentiment score of the text, resulting in the temporal sentiment feature $sen_j^i \in \mathbb{R}^{d_{sen}}$, as shown follow:

$$sen_j^i = score_j^i \cdot p_j^i \tag{2}$$

Text Vector. To represent the textual features of the microblogs, we use the pre-trained word embedding model to catch the representation of text. In this work, we use the Tencent word vector model [48] and Glove model [49] to embed the Chinese and English text, respectively. Each post is embedded to a vector sequence $[e_1, e_2, ..., e_l]$, where l is the length of the post, and $e_k \in \mathbb{R}^{d_{vec}}$ is the embedded vector obtained by the word embedding model, the d_{vec} is the dimension of the embedded vector, which is 200 in this study. To obtain the text vector, we add an meanpooling layer to catch the feature of the text. Finally, we can get the text vector $vec_j^i \in \mathbb{R}^{d_{vec}}$ as follow:

$$vec_j^i = Meanpooling(e_1, e_2, ..., e_l) \tag{3}$$

where $Meanpooling(*)$ represents the mean pooling layer, and vec_j^i represents the text vector representation of post j in event i.

Microblog Representation. After obtaining the text vector vec_j^i and the temporal sentiment feature sen_j^i, we concatenate the two vectors to obtain the microblog representation $x_j \in \mathbb{R}^{(d_{vec}+d_{sen})}$ as shown follow:

$$x_j^i = \begin{bmatrix} vec_j^i \\ sen_j^i \end{bmatrix} \tag{4}$$

3.3 Comprehensive Representation

In this study, we use RvNN to catch the feature of microblogs along the spreading path and get the comprehensive representation of the event. Furthermore, we use GRU as the hidden unit to recursively catch the features of the input microblog representation [50]. For the event C_i, the hidden state h_j^i of a node t_j^i can be computed by microblog representation x_j^i of node t_j^i and the hidden state of parent node $h_{\mathrm{P}(\mathrm{t}_\mathrm{j}^\mathrm{i})}$, where $\mathrm{P}(\mathrm{t}_\mathrm{j}^\mathrm{i})$ represents the parent node of t_j^i. Specifically, the process of RvNN can be formulated as follow:

$$h_{t_j^i} = GRU(x_j^i, h_{\mathrm{P}(t_j^i)}) \tag{5}$$

where $GRU(*)$ represents the process of GRU. As the microblog spreads, a set of the hidden states $(h_{t_q^i}, h_{t_w^i}, ..., h_{t_l^i})$ is obtained, where t_q^i, t_w^i,...,t_l^i are the leaf nodes of V_i.

After obtaining the recursive representations of all leaf nodes, we add a max-pooling layer to obtain the comprehensive representation of event C_i:

$$h_{out}^i = Maxpooling(h_{t_q^i}, h_{t_w^i}, ..., h_{t_l^i}) \tag{6}$$

where $Maxpooling(*)$ represents the max pooling layer.

3.4 Rumor Classifier

We use a multi-layer neural network with Relu activation function as a rumor classifier to claim whether the source post is a rumor. The input is the comprehensive representation h_{out}^i of event C_i, and the output $\hat{y}_i$ is the detect label of event C_i. The process of MLP shows as follow:

$$\hat{y}_i = \sigma(W_1 \cdot Relu(W_2 \cdot Relu(W_3 \cdot h_{out}^i + b_3) + b_2) + b_1) \tag{7}$$

where $\hat{y}$ denotes prediction value, W_1, W_2, W_3, b_1, b_2 and b_3 denote the weight and bias of the MLP model, $Relu(*)$ is the Relu activation function and the $\sigma(*)$ is the Sigmoid activation function. Furthermore, the cross-entropy function is adopted as the loss function, as shown as follows:

$$\mathcal{L}_\Theta(y_i, \hat{y}_i) = -y_i * \log(\hat{y}_i) + (1 - y_i) * \log(1 - \hat{y}_i) \tag{8}$$

where y_i represents the label of sample C_i and the $\mathcal{L}_\Theta(y_i, \hat{y}_i)$ represents the loss of y_i and $\hat{y}_i$.

4 Experiments

4.1 Datasets

This work employs two public datasets, i.e., Weibo [51] and RumourEval-2019 [52], to evaluate our proposed and baseline methods. The Weibo dataset has 4609 events, with rumor and non-rumor labels. The dataset includes microblog content, social content and spatiotemporal information from Weibo. It should be noted that although the original Weibo dataset has 4664 events, our study focuses on the events which have no more than about 10000 microblogs. Thus, we only remain the 4609 events for the Weibo dataset.

The RumourEval-2019 dataset has 446 events, with rumor, non-rumor and unverified labels. The dataset includes microblog content, social content and spatiotemporal information from Twitter and Reddit. Since our goal is to identify the authenticity of rumors, thus, only the rumor and non-rumor labels remain (totaling 323 events). The statistical information of the two datasets is shown in Table 1.

Table 1. Statistics of the datasets.

Feature	Weibo	RumourEval-2019
Num of microblogs	2,002,060	6,085
Num of events	4,609	323
Num of rumors	2,274	138
Num of non-rumors	2,335	185

4.2 Baseline Models

This work compares six baseline models with our proposed model. The brief introduction to the baseline model is as follows:

- **BERT** [12]**:** BERT is a pre-trained language representation model with powerful performance.
- **BiGRU** [29]**:** BiGRU uses a bidirectional-GRU to catch the feature of content words in the microblogs and can detect rumors effectively.
- **Emotion Enhanced BiGRU (Emo-BiGRU)** [29]**:** The Emo-BiGRU is the BiGRU enhanced by the dual emotion features.
- **RNN** [51]**:** This method model the social context information of events as time series of variable length and classify the events with RNN.
- **RvNN** [25]**:** The RvNN uses the TF-IDF to represent the text content and use the RvNN to catch the feature of text content alone the propagation path.
- **BiGCN** [15]**:** Bi-GCN is a bi-directional graph model and uses the top-down GCN and the bottom-up GCN to catch the features of the spreading structures of microblogs.

Where the BERT, BiGRU, and Emo-BiGRU utilize the source posts and the remaining baseline methods (RNN, RvNN, and BiGCN) utilize both source posts and replies.

4.3 Experimental Settings

Two datasets are divided into the training set, validation set, and test set, where the ratios of these three sets in the two datasets are 3:1:1 in the Weibo dataset and 7:1:2 (remain the original division ratio) in the RumourEval-2019 dataset, respectively. Furthermore, our proposed model uses the Adam optimizer [53] with a learning rate of 0.005 and sets the hidden layer dimension as 128. For the BERT model, it uses Chinese BERT pre-trained with whole word masking [54] for the Chinese text in the Weibo dataset and the google pre-trained model [12] for the English text in the RumourEval-2019 dataset. Moreover, the BERT model use an MLP with Relu activation function to classify the events, where the dimension of the hidden layer is set as 256. For the other baseline models, we adopt the original parameters.

4.4 Evaluation Metrics

Our work commonly uses the accuracy, precision, recall, and macro-F1 score as the evaluation metrics to evaluate the model's performance. The details of the evaluation metrics are as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{12}$$

$$macro - F1 = \frac{F1_{rumor} + F1_{non-rumor}}{2} \tag{13}$$

where TP represents the number of true positives, TN represents the number of false negatives, FP represents the number of false positives, FN represents the number of false negatives, $F1_{rumor}$ represents the $F1$ score of rumor and $F1_{non-rumor}$ represents the $F1$ score of non-rumor.

4.5 Experimental Results

The experimental results of our proposed method and the baseline methods are shown in Table 2 and 3. Table 2 shows the results on the Weibo dataset. The results show that our proposed model achieves the best performance compared to the baseline models. Furthermore, the poorest performance of the RNN model may be because the RNN model simplifies the propagation structures and can not catch enough features from the content of the posts. Moreover, the Emo-BiGRU's better performance than the BiGRU model also shows the effectiveness of sentiment features. The excellent performance of the BERT model in capturing text representations makes the BERT model perform better than Emo-BiGRU and BiGRU. Finally, compared to the BERT, the better performance of RvNN and BiGCN shows the importance of the spreading structures in rumor detection, since the spreading structures are hard to be manipulated by malicious users.

Table 3 shows the experimental results on the RumourEval-2019 dataset. Compared to experiments on the Weibo dataset, the results on the RumourEval-2019 dataset are worse than that on the Weibo dataset. This may be caused by the sparse data and the low inter-annotator agreement of labels of the RumourEval-2019 dataset [55,56]. In Ref. [55] and Ref. [56], the authors pointed out that the rate of overall inter-annotator agreement is 63.7% which means that there are many conflicting or inconsistent labels and leading to worse performance.

Table 2. Rumor detection performance on the Weibo dataset.

Method	*Accuracy*	*macro-F1*	Rumor			Non-rumor		
			Precision	*Recall*	*F1*	*Precision*	*Recall*	*F1*
BERT	0.911	0.912	0.919	0.900	0.910	0.904	0.922	0.913
BiGRU	0.793	0.792	0.810	0.758	0.783	0.778	0.826	0.801
Emo-BiGRU	0.857	0.857	0.847	0.866	0.857	0.866	0.848	0.857
RNN	0.630	0.629	0.611	0.690	0.648	0.654	0.571	0.610
RvNN	0.919	0.919	0.938	0.895	0.916	0.902	0.942	0.921
BiGCN	0.921	0.677	0.675	0.690	0.676	0.679	0.689	0.678
Ours	**0.939**	**0.939**	**0.944**	**0.925**	**0.934**	**0.935**	**0.951**	**0.943**

Table 3. Rumor detection performance on the RumourEval-2019 dataset.

Method	*Accuracy*	*macro-F1*	Rumor			Non-rumor		
			Precision	*Recall*	*F1*	*Precision*	*Recall*	*F1*
BERT	0.509	0.526	**0.496**	0.475	0.478	0.565	0.594	0.573
BiGRU	0.423	0.421	0.486	0.425	0.453	0.361	0.419	0.388
Emo-BiGRU	0.423	0.423	0.484	0.375	0.423	0.375	0.484	0.423
RNN	0.408	0.282	0.418	**0.903**	0.517	0.250	0.025	0.046
RvNN	0.507	0.506	0.455	0.645	**0.533**	**0.593**	0.400	0.478
BiGCN	**0.600**	0.424	0.371	0.400	0.371	0.486	0.514	0.476
Ours	0.549	**0.534**	0.481	0.419	0.448	0.591	**0.650**	**0.619**

Nevertheless, our proposed approach still achieves the best macro-F1. Moreover, the performance of BiGCN demonstrates the robustness of the propagation structure features on different datasets, which helps the BiGCN get the best accuracy. Furthermore, the observed phenomenon of the better result in detecting the rumors and the worst result in detecting the non-rumors of RNN implies RNN is overfitting with limited data. Moreover, the Emo-BiGRU also can be observed that have a better result than BiGRU in macro-F1. The RvNN can use both the text content and the propagation structure, which helps RvNN perform better than Emo-BiGRU. The performance of BERT shows that BERT can effectively catch the features of the text content.

4.6 Discussions

Ablation Study. To further investigate the effectiveness of the key components of our proposed model, i.e., the temporal sentiment feature and the text vector, we additionally conduct an ablation study. In particular, we consider two types of ablations in our experiments: 1) *Ours w/o Text* that does not generate text vector in microblog representation, 2) *Ours w/o Sen* that does not generate the temporal sentiment feature in microblog representation. As the results are shown in Table 4, we observe that *Ours w/o Text* gets the poorest performance in the

Weibo dataset, implying the textual feature is more important than our temporal sentiment feature in detective the rumors in the Weibo dataset. Furthermore, *Ours w/o Sen* has worse performance than *Ours w/o Text* in the RumourEval-2019 dataset implies that the temporal sentiment feature is more important than the textual feature in the RumourEval-2019 dataset. Moreover, our proposed model with all key components has the best performance in both two datasets demonstrating the validity of the proposed microblog representation. The results also demonstrate that our temporal sentiment feature plays an important role in the sparse dataset, especially those with shallower reply depths, e.g., the RumourEval-2019 dataset.

Table 4. The results of ablation study.

Dataset	Metrics	w/o Text	w/o Sen	Ours
Weibo	*Accuracy*	0.855	0.926	**0.939**
	Precision	0.855	0.925	**0.939**
	Recall	0.854	0.925	**0.938**
	*macro-F*1	0.855	0.925	**0.939**
RumourEval-2019	*Accuracy*	0.493	0.493	**0.549**
	Precision	0.511	0.496	**0.536**
	Recall	0.510	0.496	**0.535**
	*macro-F*1	0.490	0.492	**0.534**

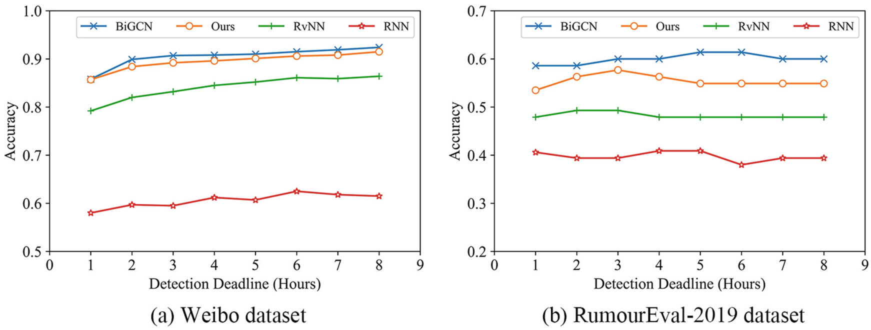
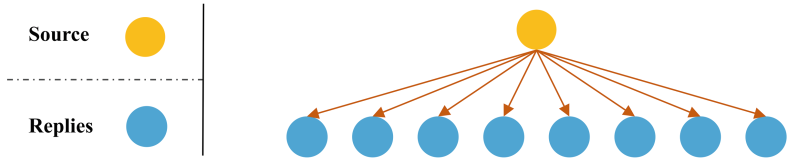


Fig. 2. Early rumor detection accuracy of different methods on two datasets.

Early Rumor Detection. Early rumor detection is an important metric for evaluating the quality of the method in the early stage of information spread. Since the BERT, BiGRU and Emo-BiGRU mainly utilize the source post, early detection is meaningless for these methods. In this work, we compare our proposed method with the RNN, RvNN, and BiGCN in the Weibo and RumourEval-2019 datasets, as shown in Fig. 2. From Fig. 2, we can find that our proposed method reaches high accuracy at the early stage of the propagation. However, our model still performs worse than the BiGCN, which may be because the sentiment features do not appear significant in the early state of rumor spreading, especially for the situation when the propagation features are significant. We have experimented with the effective time of the temporal sentiment feature, and it shows that the temporal sentiment feature requires at least 17 h after source microblogs are posted.



Layer	Text
0	Category 6? If Hurricane Irma Becomes The Strongest Hurricane In History, It Could Wipe Entire Cities Off The Map
1	I dont think 6 is possible but stay safe my Florida Friends and all others in IRMA's Path!"
1	@[redacted] Photos of that hurricane make it look even more of a monstrosity than satellite imagery of it! Hope that
1	@[redacted] Watch what the power of Prayer does
1	@[redacted] @[redacted] Pray to god the only chance you have
1	@[redacted] @[redacted] Dropping now, is Cat 3 Wed.
1	@[redacted] @[redacted] no such thing as category 6, but stay prepared, it's a dangerous storm and track is still unpredictable.
1	@[redacted] @[redacted] Get out of there people. Florida, South Carolina. Check on neighbors, give them a way to
1	@[redacted] @[redacted] I've never heard of a H6! No danger of it getting stronger as it crosses land, Although, it co

Fig. 3. A sample of false rumor in RumourEval-2019 dataset.

Case Study. To show the importance of the temporal sentiment feature, we demonstrate a case study in Fig. 3. In this case study, the microblog only contains the source post and one layer of replies, which implies there have not enough propagation structure features. Furthermore, the text content also shows there lack the direct evidence to judge whether it is a rumor, such as "false" or "fake", which leads to the *w/o Text* model misjudging this case as true. However, our proposed model with the temporal sentiment feature can correctly identify this case as a rumor, suggesting that the sentiment features do play an important role in the absence of textual and communication features.

5 Conclusion

Since the sentiment is a hidden feature and plays an important role in rumor spreading, this work proposes a rumor detection model based on temporal sentiment features. The experiments on two real datasets (the Weibo and RumourEval-2019) demonstrate that our proposed model has better performance than the baseline models, and the ablation study shows that our temporal sentiment feature is effective for rumor detection. However, our method performs worse than Bi-GCN in the early rumor detection. This may be because the temporal sentiment features focus on the long-term sentiment change as information spreads. In our future work, we will consider the importance of different replies to enhance early rumor detection performance. In conclusion, our work supplements current rumor detection methods and highlights that the proposed temporal sentiment feature can effectively capture the rumor in the propagation, especially when there are not enough propagation and text features.

Acknowledgements. This work was supported by the Zhejiang Fundamental Public Welfare Research Project under Grant (LGF21G010003 and LGF20F020016) and the National Natural Science Foundation of China under Grant (62103374).

References

1. Shao, C., Ciampaglia, G.L., Varol, O., Yang, K.C., Flammini, A., Menczer, F.: The spread of low-credibility content by social bots. Nature Commun. **9**(1), 4787 (2018)
2. Ruan, Z., Yu, B., Shu, X., Zhang, Q., Xuan, Q.: The impact of malicious nodes on the spreading of false information. Chaos: Interdisciplinary J. Nonlinear Sci. **30**(8), 083101 (2020)
3. Pendleton, S.C.: Rumor research revisited and expanded. Lang. Commun. **18**(1), 69–86 (1998)
4. Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., Procter, R.: Detection and resolution of rumours in social media: a survey. ACM Comput. Surv. **51**(2), 1–36 (2018)
5. Allcott, H., Gentzkow, M.: Social media and fake news in the 2016 election. J. Econ. Perspect. **31**(2), 211–36 (2017). May
6. Du, J., Dou, Y., Xia, C., Cui, L., Ma, J., Philip, S.Y.: Cross-lingual covid-19 fake news detection. In: Proceedings of the 16th International Conference on Data Mining Workshops, pp. 859–862. IEEE Computer Society, Los Alamitos (2021)
7. Cui, L., Lee, D.: CoAID: COVID-19 healthcare misinformation dataset. arXiv preprint. arXiv:2006.00885 (2020)
8. Hassan, N., Li, C., Tremayne, M.: Detecting check-worthy factual claims in presidential debates. In: Proceedings of the 24th International on Conference on Information and Knowledge Management, pp. 1835–1838. ACM, Atlanta (2015)
9. Rubin, V.L.: On deception and deception detection: content analysis of computer-mediated stated beliefs. In: Proceedings of the 73rd Annual Meeting on Navigating Streams in an Information Ecosystem, vol. 47, p. 10. ASIS, Pittsburgh (2010)
10. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint. arXiv:1301.3781 (2013)
11. Vaswani, A.: Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, pp. 6000–6010. CAI, Long Beach (2017)

12. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. arXiv preprint. arXiv:1810.04805 (2018)
13. Chenbo, F., Zheng, Y., Liu, Y., Xuan, Q., Chen, G.: NES-TL: network embedding similarity-based transfer learning. IEEE Trans. Netw. Sci. Eng. **7**(3), 1607–1618 (2020)
14. Chenbo, F., et al.: A novel spatiotemporal behavior-enabled random walk strategy on online social platforms. IEEE Trans. Comput. Soc. Syst. **9**(3), 807–817 (2022)
15. Bian, T.: Rumor detection on social media with bi-directional graph convolutional networks. arXiv preprint. arXiv:2001.06362 (2020)
16. Kaliyar, R.K., Goswami, A., Narang, P.: FakeBERT: fake news detection in social media with a bert-based deep learning approach. Multimedia Tools Appl. **80**(8), 11765–11788 (2021). https://doi.org/10.1007/s11042-020-10183-2
17. Qian, S., Wang, J., Hu, J., Fang, Q., Xu, C.: Hierarchical multi-modal contextual attention network for fake news detection. In: Proceedings of the 44th International Conference on Research and Development in Information Retrieval, pp. 153–162. ACM, New York (2021)
18. Qazi, M., Khan, M.U., Ali, M.: Detection of fake news using transformer model. In: Proceedings of the 3rd International Conference on Computing, Mathematics and Engineering Technologies, pp. 1–6. IEEE, Sukkur (2020)
19. Lu, Y.J., Li, C.T.: GCAN: graph-aware co-attention networks for explainable fake news detection on social media. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 505–514. ACL (2020)
20. Feng, S., Banerjee, R., Choi, Y.: Syntactic stylometry for deception detection. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, vol. 2, pp. 171–175. ACL, Jeju Island (2012)
21. Wang, Y.: EANN: event adversarial neural networks for multi-modal fake news detection. In: Proceedings of the 24th International Conference on Knowledge Discovery & Data Mining, pp. 849–857. ACM, East Lansing (2018)
22. Thies, J., Zollhöfer, M., Stamminger, M., Theobalt, C., Nießner, M.: Face2face: Real-time face capture and reenactment of RGB videos. Commun. ACM **62**(1), 96–104 (2018). Dec
23. Zhou, Z., Guan, H., Bhat, M.M., Hsu, J.: Fake news detection via nlp is vulnerable to adversarial attacks. arXiv preprint. arXiv:1901.09657 (2019)
24. Westerlund, M.: The emergence of deepfake technology: a review. Technol. Innov. Manag. Rev. **9**, 40–53 (2019)
25. Ma, J., Gao, W., Wong, K.F.: Rumor detection on Twitter with tree-structured recursive neural networks. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, vol. 1, pp. 1980–1989. ACL, Melbourne (2018)
26. Wu, K., Yang, S., Zhu, K.Q.: False rumors detection on sina weibo by propagation structures. In: Proceedings of the 31st International Conference on Data Engineering, pp. 651–662. IEEE Computer Society, Los Alamitos (2015)
27. Zhiyuan, W., Pi, D., Chen, J., Xie, M., Cao, J.: Rumor detection based on propagation graph neural network with attention mechanism. Expert Syst. Appl. **158**, 113595 (2020)
28. Alonso, M.A., Vilares, D., Gómez-Rodríguez, C., Vilares, J.: Sentiment analysis for fake news detection. Electronics **10**(11), 1348 (2021)
29. Zhang, X., Cao, J., Li, X., Sheng, Q., Zhong, L., Shu, K.: Mining dual emotion for fake news detection. In: Proceedings of the 30th the Web Conference, pp. 3465–3476. ACM, Ljubljana (2021)

30. Mackey, A., Gauch, S., Labille, K.: Detecting fake news through emotion analysis. In: Proceedings of the 13th International Conference on Information, Process, and Knowledge Management, pp. 65–71. IARIA (2021)
31. Yang, Y., Zheng, L., Zhang, J., Cui, Q., Li, Z., Yu, P.S.: TI-CNN: convolutional neural networks for fake news detection. arXiv preprint. arXiv:1806.00749 (2018)
32. Wang, Z., Guo, Y., Wang, J., Li, Z., Tang, M.: Rumor events detection from chinese microblogs via sentiments enhancement. IEEE Access **7**, 103000–103018 (2019)
33. Ajao, O., Bhowmik, D., Zargari, S.: Sentiment aware fake news detection on online social networks. In: Proceedings of the 44th International Conference on Acoustics, Speech, and Signal Processing, pp. 2507–2511. IEEE, Brighton (2019)
34. Cui, L., Wang, S., Lee, D.: SAME: sentiment-aware multi-modal embedding for detecting fake news. In: Proceedings of the 4th International Conference on Advances in Social Networks Analysis and Mining, pp. 41–48. ACM, Vancouver (2019)
35. Wu, L., Rao, Y.: Adaptive interaction fusion networks for fake news detection. arXiv preprint. arXiv:2004.10009 (2020)
36. Davoudi, M., Moosavi, M.R., Sadreddini, M.H.: DSS: a hybrid deep model for fake news detection using propagation tree and stance network. Expert Syst. Appl. **198**, 116635 (2022)
37. Bond, G.: 'Lyin' Ted', 'Crooked Hillary', and 'Deceptive Donald': language of lies in the 2016 US presidential debates. Appl. Cogn. Psychol. **31**(6), 668–677 (2017)
38. Zhao, H., Zhou, W., Chen, D., Wei, T., Zhang, W., Yu, N.: Multi-attentional deepfake detection. In Proceedings of the 30st Conference on Computer Vision and Pattern Recognition, pp. 2185–2194. IEEE Computer Society, Los Alamitos (2021)
39. Guarnera, L., Giudice, O., Battiato, S.: Deepfake detection by analyzing convolutional traces. In: Proceedings of the 29th Conference on Computer Vision and Pattern Recognition Workshops, pp. 2841–2850. IEEE Computer Society, Los Alamitos (2020)
40. Ma, J., Gao, W., Wong, K.-F.: Detect rumors in microblog posts using propagation structure via kernel learning. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, vol. 1, pp. 708–717. ACL, Vancouver (2017)
41. Li, B., Qian, Z., Li, P., Zhu, Q.: Multi-modal fusion network for rumor detection with texts and images. In: Þór Jónsson, B., et al. (eds.) MMM 2022. LNCS, vol. 13141, pp. 15–27. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-98358-1_2
42. Vo, N., Lee, K.: Hierarchical multi-head attentive network for evidence-aware fake news detection. arXiv preprint. arXiv:2102.02680 (2021)
43. Dou, Y., Shu, K., Xia, C., Yu, P.S., Sun, L.: User preference-aware fake news detection. In: Proceedings of the 44th International Conference on Research and Development in Information Retrieval, pp. 2051–2055. ACM, Montreal (2021)
44. Wang, G., Tan, L., Shang, Z., Liu, H.: Multimodal dual emotion with fusion of visual sentiment for rumor detection. arXiv preprint. arXiv:2204.11515 (2022)
45. Guo, C., Cao, J., Zhang, X., Shu, K., Yu, M.: Exploiting emotions for fake news detection on social media. arXiv preprint. arXiv:1903.01728 (2019)
46. Guerini, M., Staiano, J.: Deep feelings: a massive cross-lingual study on the relation between emotions and virality. In: Proceedings of the 24th International Conference on World Wide Web, pp. 299–305. Association for Computing Machinery, New York (2015)

47. Pröllochs, N., Bär, D., Feuerriegel, S.: Emotions explain differences in the diffusion of true vs. false social media rumors. Sci. Rep. **11**(1), 1–12 (2021)
48. Song, Y., Shi, S., Li, J., Zhang, H.: Directional KKIP-Gram: explicitly distinguishing left and right context for word embeddings. In: Proceedings of the 16th Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vol. 2, pp. 175–180. ACL, New Orleans (2018)
49. Pennington, J., Socher, R., Manning, C.: GloVe: global vectors for word representation. In: Proceedings of the 19th Conference on Empirical Methods in Natural Language Processing, pp. 1532–1543. ACL, Doha (2014)
50. Cho, K.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint. arXiv:1406.1078 (2014)
51. Ma, J.: Detecting rumors from microblogs with recurrent neural networks. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence, pp. 3818–3824. AAAI Press, New York (2016)
52. Gorrell, G., Bontcheva, K., Derczynski, L., Kochkina, E., Liakata, M., Zubiaga, A.: Rumoureval 2019: determining rumour veracity and support for rumours. arXiv preprint. arXiv:1809.06683 (2018)
53. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint. arXiv:1412.6980 (2014)
54. Cui, Y., Che, W., Liu, T., Qin, B., Yang, Z.: Pre-training with whole word masking for chinese bert. IEEE/ACM Trans. Audio Speech Lang. Process. **29**, 3504–3514 (2021)
55. Gorrell, G., et al.: SemEval-2019 task 7: RumourEval, determining rumour veracity and support for rumours. In: Proceedings of the 13th International Workshop on Semantic Evaluation, pp. 845–854. ACL, Minneapolis (2019)
56. Li, Q., Zhang, Q., Si, L.: eventAI at SemEval-2019 task 7: rumor detection on social media by exploiting content, user credibility and propagation information. In: Proceedings of the 13th International Workshop on Semantic Evaluation, pp. 855–859. ACL, Minneapolis (2019)