

## *Retraction*

# **Retracted: A Multichannel Model for Microbial Key Event Extraction Based on Feature Fusion and Attention Mechanism**

## **Security and Communication Networks**

Received 26 December 2023; Accepted 26 December 2023; Published 29 December 2023

Copyright © 2023 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

## **References**

- [1] P. Li and Q. Wang, "A Multichannel Model for Microbial Key Event Extraction Based on Feature Fusion and Attention Mechanism," *Security and Communication Networks*, vol. 2021, Article ID 7800144, 10 pages, 2021.

## Research Article

# A Multichannel Model for Microbial Key Event Extraction Based on Feature Fusion and Attention Mechanism

Peng Li<sup>1</sup> and Qian Wang<sup>2</sup>

<sup>1</sup>Teaching and Research Department of Ideological and Political Theory, Kaifeng University, Kaifeng 475004, Henan, China

<sup>2</sup>College of Humanities, Kaifeng University, Kaifeng 475004, Henan, China

Correspondence should be addressed to Peng Li; 205912095@kfu.edu.cn

Received 9 November 2021; Revised 9 December 2021; Accepted 18 December 2021; Published 31 December 2021

Academic Editor: Jian Su

Copyright © 2021 Peng Li and Qian Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to further mine the deep semantic information of the microbial text of public health emergencies, this paper proposes a multichannel microbial sentiment analysis model MCMF-A. Firstly, we use word2vec and fastText to generate word vectors in the feature vector embedding layer and fuse them with lexical and location feature vectors; secondly, we build a multichannel layer based on CNN and BiLSTM to extract local and global features of the microbial text; then we build an attention mechanism layer to extract the important semantic features of the microbial text; thirdly, we merge the multichannel output in the fusion layer and use soft; finally, the results are merged in the fusion layer, and a surtax function is used in the output layer for sentiment classification. The results show that the F1 value of the MCMF-A sentiment analysis model reaches 90.21%, which is 9.71% and 9.14% higher than the benchmark CNN and BiLSTM models, respectively. The constructed dataset is small in size, and the multimodal information such as images and speech has not been considered.

## 1. Introduction

The open and social nature of social networking platforms has led to their rapid development, which corresponds to the explosive growth of microbiology data [1]. Social events on Weibo are transmitted through social networks or news websites, and social hotspots are generated through fermentation. As time passes and events develop, hotspots evolve dynamically, generating different key information at each timestamp, which contains intricate developmental and evolutionary relationships among events. Although Weibo has a rich resource of social hot events, it is difficult for the users to capture key information at each evolutionary stage of social hot events in the face of massive data [2]. As a representative platform of social networks, Sina Weibo has made great contributions to the dissemination of events, but from the perspective of social network information dissemination, the “retreating” feature of Weibo brings a lot of redundant information and leads to a flood of information. Therefore, the filtering of social hotspots and the extraction

of key events are important for users to understand the social hotspots and track the evolution of hotspots [3–5].

In addition, extracting key events from social hotspots is also of research significance for decision-makers to analyze the public opinion situation and guide social opinions.

At present, the research work related to social hot events is based on the content characteristics of events, and the traditional method of term frequency-inverse document frequency (TF-IDF) model is the most commonly used model to measure the importance of texts, but the TF-IDF model can only reveal the importance of events at the word level [6]. Recently, some researchers have analyzed online opinion events based on Bayesian networks [7], and some studies have proposed to model events as graph structures and use dominating set algorithms or calculate the degree and aggregation coefficients of graph nodes to extract key events by considering the importance of nodes from the perspective of graph theory [8], but constructing graphs requires rich semantic information of events, which is lacking in the microbial data. In conclusion, all the above

studies ignore the influence of the propagation of events in the microbiology environment on the key events.

To solve the above problems, this paper proposes a microbial key event extraction method that integrates the social influence and temporal distribution of events. Firstly, we model the social influence of events based on the features of events in microbiology; then we extract key events under different time distributions based on the features of the temporal distribution of event evolution; finally, experiments based on real microbial datasets show that the method in this paper can effectively extract key events in each evolutionary stage of social hot events. The main work of this paper is as follows:

A method for modeling the importance of microbiology events is proposed by building a social influence model related to the event topic and mining the important elements of microbial events.

Establishing a microbial key event detection model that incorporates the social influence of events and time distribution.

The proposed extraction method is experimentally validated on two real microbial datasets and a microbial key event extraction system is constructed.

## 2. Related Work

Data mining for social networks is an important research direction in web text mining. The data mining analysis for microbiology generally includes topical event mining, sentiment analysis, and web opinion analysis. Among them, topic event mining includes high-quality information extraction [9] and event evolution mining [10], while the current high-quality information extraction focuses on event extraction [11] and event summary [12], which is different from the key event extraction research that the work of this paper focuses on. Current key event extraction methods in social hotspots can be divided into traditional content feature-based methods, graph-based methods, and machine learning-based methods, and the characteristics and shortcomings of different methods are briefly described below:

- (1) Traditional content feature-based methods. This type of method uses event content features to evaluate and rank events and extracts key events in score ranking. For example, the authors of [13] calculated the TF-IDF score of keywords in the events, uses the sum of the keyword score as the event score and extracts the events with the top score as key events. The authors of [14] fused multiple features of events and transforms them to wavelet domain to capture the detailed differences between events and introduces kernel principal component analysis for feature transformation to extract key events. The authors of [15] used event hotspots to divide opinion relations, extracts event keywords and text summaries of key events based on the textrank algorithm [16], and finally builds a master map and visualizes event summaries.

- (2) Graph-based methods. This class of methods shows events as graph structures based on event content feature relations and uses graph algorithms to transform extracting key events into extracting key nodes in the graph. For example, [17] models microbiology as multiview graphs based on similarity, solves important nodes to extract key events using minimum weight dominating sets, and introduces TopK sets to alleviate the problem of a huge volume of microbial data. And the authors of [18] introduced degree and aggregation coefficients [19] to evaluate the importance of nodes in the graph to extract key events.
- (3) Machine learning-based methods. This type of method uses machine learning algorithms to model and learn how events achieve key event extraction. For example, the authors of [20] fused network representation learning with a K-means clustering algorithm to represent low-dimensional vectors of public opinion events and cluster them to obtain public opinion events. In addition, the authors of [21] used the K-mean clustering algorithm, K-nearest neighbor classification algorithm, and decision tree three-class method to model the geographical features of microbiology events and detect and extract key events from different geographical locations.

The above methods evaluate the event importance from the perspective of event content features and introduce external models for algorithm enhancement, but ignore the impact of event propagation on event importance, and although the authors of [22] introduced the propagation behavior features of microblog events, the complex mathematical transformation will lead to huge time overhead. Although the machine learning-based approach models event features at a fine-grained level, it ignores the characteristics of events in social networks. In this paper, we use the social influence of events to make up for the shortage of content feature-based methods, and in addition, we introduce the event time distribution to maximize the reasonableness of the distribution of extracted events on the timeline and improve the accuracy of key event extraction.

## 3. Experimental Procedure and Model Construction

*3.1. Experimental Procedure.* The experimental process is shown in Figure 1. Where we first obtain the microbial data through crawlers and perform preprocessing and lexical annotation. Secondly, word2vec and fastText are used to generate word vectors from the microbiology and fuse them with lexical and location features to fully explore the subtle sentiment information contained in the sentiment words and location relationships in the microbiology; then, the multifeature fusion vectors are fed into a multichannel layer constructed by the CNN and BiLSTM, and an attention mechanism is introduced afterward; then a multichannel microbial sentiment analysis model MCMF-A is proposed

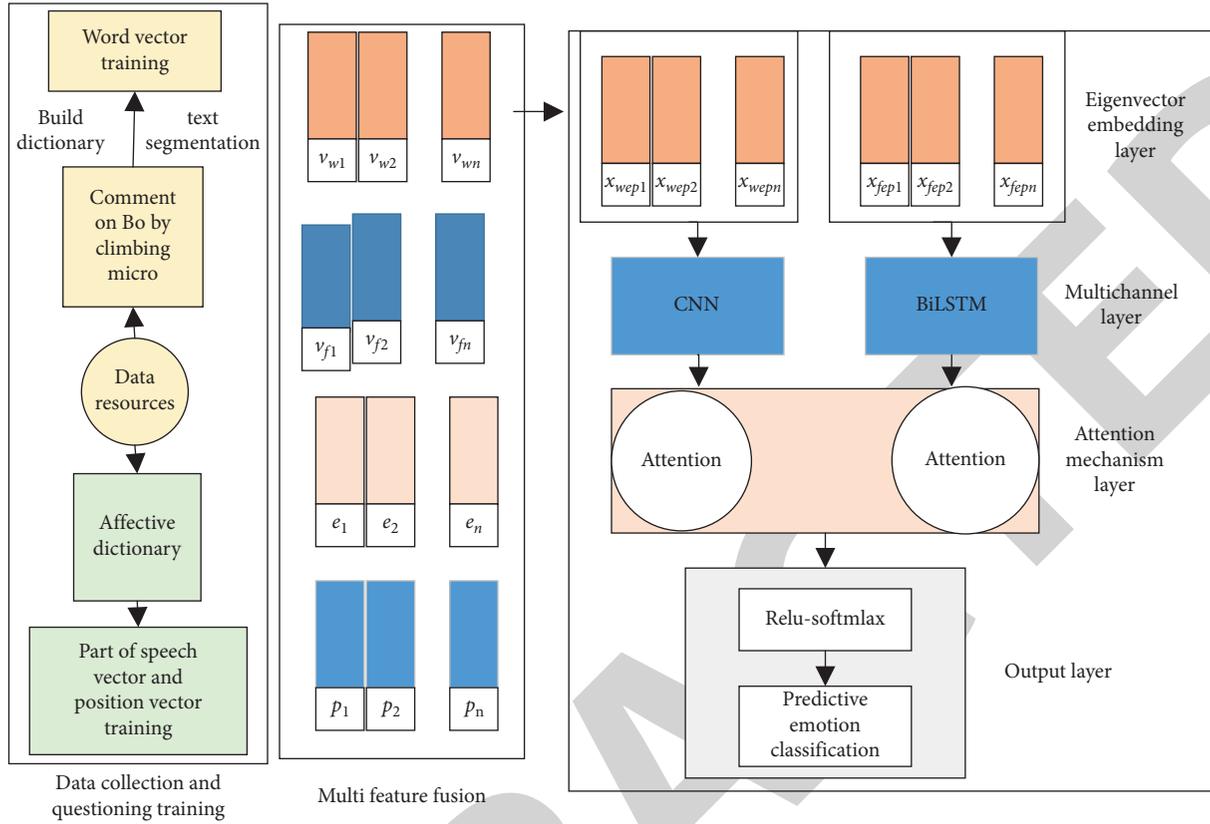


FIGURE 1: Experimental flow chart.

based on feature fusion and attention mechanism. Finally, we propose MCMF-A, a multichannel microbial sentiment analysis model based on feature fusion and attention mechanism, which can more fully utilize the global and local deep semantic feature information of the microbial text and further improve the microbial sentiment analysis [22].

**3.2. Microblog Sentiment Analysis Model Construction.** Based on CNN and BiLSTM, this paper proposes a multichannel microbial sentiment analysis model MCMF-A based on feature fusion and attention mechanism, which are a five-layer structure, as shown in Figure 1.

Word2vec and fastText is used to generate word vectors, which are fused with part of speech feature (POSF) generated from lexical information and position feature (PF) generated from location information to generate multi-feature fusion vectors WMF and FMF. Fusing the word2vec and fastText word vectors with POSF generated from lexical information and PF generated from location information to generate the multifeature fusion vectors WMF and FMF [23].

The WMF and FMF are used as input to construct four channels, and CNN and BiLSTM are used to extract the local features, respectively, and BiLSTM to extract local semantic features and global semantic features, respectively.

The attention mechanism is introduced after the multichannel hybrid neural network model to fully consider the important features in the microbiology text.

The output features of the fusion layer are used to predict the sentiment classification by the surtax function.

### 3.3. Feature Vector Embedding Layer

**3.3.1. Lexical Feature Vector.** Due to the limitation of the microbial growth and the existence of a large number of short texts and colloquial expressions, the available feature information is very limited. In order to fully exploit the lexical feature information in microbiology, this study uses HowNet sentiment dictionary to lexically annotate six types of sentiment words with the help of jieba and generates  $m$ -dimensional continuous value vector POSF by vector mapping operation for specific sentiment words appearing in the sentiment dictionary in microbiology to fully extract the subtle sentiment information in microbiology [24].

**3.3.2. Location Feature Vector.** There are interactions among emotion words, degree words, and negation words in the microbiology text, and the same word appears. The degree of emotion expressed by the same word in different positions may be different, and the position value [25] also varies. The positional value is a value that indicates the relationship between words. The position value is a numerical value that indicates the relationship between words and generates PF to better represent the importance of words in a sentence. The positional value is calculated as in (1):

$$\text{position}_i = \begin{cases} e_i = 0, & \text{MAX}_{SENT\_LEN} - \text{len}(\text{words}) + i, \\ e_i \neq 0, & \text{MAX}_{SENT\_LEN} + i, \end{cases} \quad (1)$$

where  $\text{position}_i$  denotes the position value of the  $i^{\text{th}}$  word in the text sequence;  $e_i \neq 0$  denotes that the word contains lexical features;  $\text{MAX}_{SENT\_LEN}$  is the maximum step size of the input text sequence;  $\text{len}(\text{words})$  is the length of the text sequence. Similar to the generating POSF method, this study generates  $r$ -dimensional continuous value vector PF for each location value by vector mapping operation in the location feature vector representation, the  $i^{\text{th}}$  word PF is defined as  $p_i \in R^r$ .

**3.3.3. Multifeature Fusion Vector.** In order to fully consider the semantic relationship between features [26], the word vectors generated by word2vec and fastText are fused with POSF and PF to generate multifeature fusion vectors WMF and FMF. The fusion calculation is shown in equations (2) and (3), where  $\oplus$  is the fusion operation and  $n$  is the length of the text sequence. The specific process is shown in Figure 2. First, the  $i^{\text{th}}$  word generates word vector matrices  $v_{wi}$  and  $v_{fi}$  with dimension size through word2vec and fastText; then, it is fused with parts of speech feature vector  $e_i$  and position feature vector  $p_i$  to generate multifeature fusion vector matrices  $x_{wep_i}$  and  $x_{fep_i}$  with dimension size of  $d + m + r$ ; finally, the word-level features are spliced into sentence-level features as the input of the next layer.

$$x_{wep_i} = v_{wi} \oplus e_i \oplus p_i, \quad x_{wep_i} \in R^{d+m+r}, \quad (2)$$

$$X_{wep} = [x_{wep_1}, x_{wep_2}, \dots, x_{wep_n}].$$

$$x_{fep_i} = v_{fi} \oplus e_i \oplus p_i, \quad x_{fep_i} \in R^{d+m+r}, \quad (3)$$

$$X_{fep} = [x_{fep_1}, x_{fep_2}, \dots, x_{fep_n}].$$

**3.4. Multichannel Layer.** The multichannel layer allows for further feature extraction of the fusion vectors, where the WMF is fed into the CNN and BiLSTM, and similarly, the FMF is fed into both models, resulting in four channels to extract the local and global features from the previous layer of the model.

The structure of the WMF-based neural network is given in Figure 3, and the FMF-based structure is similar to it. In Figure 3, on the one hand, three feature vector maps are generated by convolutional operations using convolutional kernels of different sizes, and then a max-pooling strategy is applied to obtain more comprehensive local features; on the other hand, sentiment features in both positive and negative directions are extracted and spliced based on BiLSTM, and then contextual global features are obtained. The computation is as in equations (4) and (5).

$$X_{1:n} = [x_1, x_2, \dots, x_n], \quad x_i \in R^{d+m+r},$$

$$C_{h_s} = \text{relu}(W_s \cdot X_{i:i+h_s-1} + b_s), \quad (4)$$

$$\hat{C}_{h_s} = \max(C_{h_s}),$$

$$v_{h_s} = (\hat{C}_{h_{s1}}, \hat{C}_{h_{s2}}, \dots, \hat{C}_{h_{sm}}),$$

$$v_i = [v_{h_1}, v_{h_2}, v_{h_3}], \quad (5)$$

$$H_c = [v_1, v_2, \dots, v_n],$$

where  $x_i$  is the vector of the  $i^{\text{th}}$  word,  $W_s$  is the convolutional kernel,  $h_s$  is the convolutional kernel window size ( $1 \leq s \leq 3$ ),  $b_s$  is the bias term,  $m$  is the number of each convolutional kernel,  $\text{relu}$  is the activation function, and  $H_c$  is the CNN channel feature.

$$\vec{h} = \overrightarrow{\text{LSTM}}(x_i),$$

$$\overleftarrow{h} = \overleftarrow{\text{LSTM}}(x_i),$$

$$h_i = [\vec{h}, \overleftarrow{h}],$$

$$H_b = [h_1, h_2, \dots, h_n],$$

where  $\overrightarrow{\text{LSTM}}$  is the forward propagation LSTM,  $\overleftarrow{\text{LSTM}}$  is the backward propagation LSTM,  $h_i$  is the splicing result of  $\vec{h}$  and  $\overleftarrow{h}$ , and  $H_b$  is the BiLSTM channel feature.

**3.5. Attention Mechanism Layer.** In view of the better performance of the attention mechanism in machine translation, entity recognition, and sequence annotation tasks [27], this experiment introduces the attention mechanism after the multichannel layer by weighting the vectors to extract important features. Specifically, firstly, a fully connected computation is performed on the hidden layer vector  $h_i$  using the nonlinear activation function  $\tanh$  to obtain  $u_i$ ; secondly, the attention weight  $\alpha_i$  is learned for each word by normalizing it with the softmax function; finally, the weight  $\alpha$  is weighted and summed with the corresponding hidden layer vector  $h_i$  as the attention output  $H_a$ , which is calculated as in (7).

$$u_i = \tanh(W_w h_i + b_w),$$

$$\alpha_i = \text{soft max}(u_i), \quad (7)$$

$$H_a = \sum \alpha_i h_i.$$

**3.6. Fusion Layer.** The MCMF-A fusion layer can merge the output of the features from multiple channels to generate multifeature vectors, which can then exploit the multilayer semantic information of the microblogging text. Specifically, firstly, local features are obtained by the CNN iterative computation; secondly, global features are obtained by combining the semantic information with BiLSTM; finally, an attention mechanism is introduced for secondary extraction to retain the important features [28].

In the model training process, binary cross-entropy is used as the loss function, calculated as in (8), where  $i$  is the

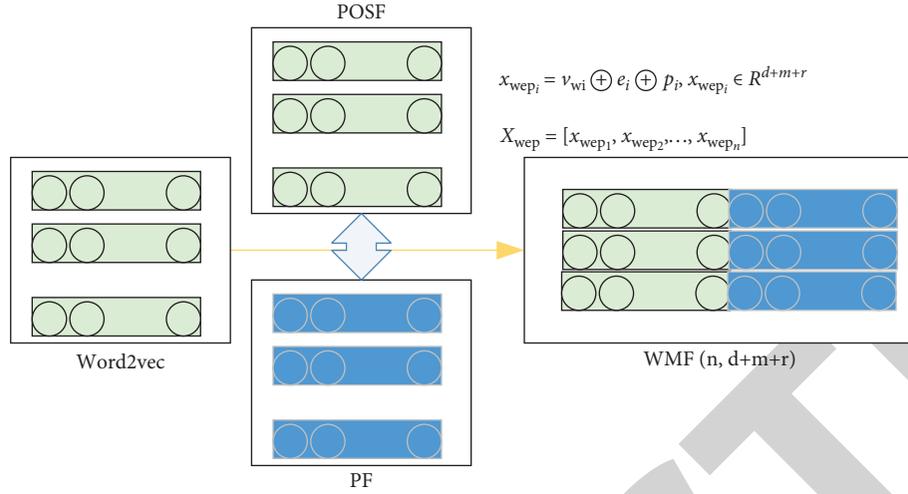


FIGURE 2: Multifeature vector fusion process.

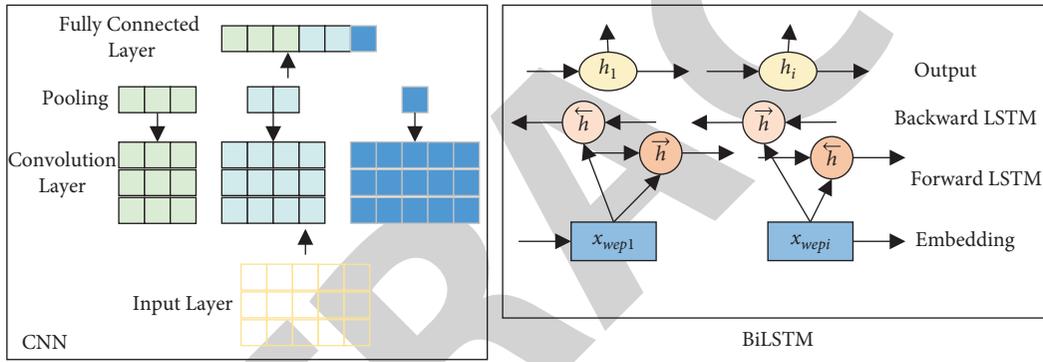


FIGURE 3: Structure of WMF-based neural network.

microblogging text sequence index,  $k$  is the sentiment category index,  $D$  is the training data set size,  $C$  is the number of categories,  $P$  is the predicted category, and  $Y$  is the actual category.

$$\text{loss} = \text{binary\_cross entropy} = - \sum_{i=1}^D \sum_{k=1}^C y_i^k \log(p_i^k). \quad (8)$$

**3.7. Output Layer.** Based on the feature vectors extracted by the fusion layer, the output layer uses the dropout mechanism to randomly select neural units as dropout elements, and the obtained results are combined with softmax [29] to perform sentiment classification. The classification prediction is calculated as in (9), where  $H$  is the feature vector extracted by the fusion layer,  $Y$  is the predicted category.

$$Y = \arg \max(\text{soft max}(WH + b)). \quad (9)$$

## 4. Experimental Results and Analysis

**4.1. Dataset.** The experiment used 5 million microbiology text provided by the NLPPIR as the experimental dataset [30]. The dataset contained necessary information such as

microbial ID, user ID, posting time, body content, number of retreats, likes, and comments. All data from March 1–31, 2014 were used as the test data set.

Figure 4 shows the distribution and percentage of tweets containing emotional symbols in terms of days. It can be seen that the number of tweets containing emotion symbols change with the total number of tweets per day, and the number of tweets containing emotion symbols account for 36% of the total on average, which also reflect the feasibility of using emotion symbols to construct emotional features to detect unexpected events.

The lines in Figure 5 indicate the frequency of emotion symbols in the data set. Figure 5 shows the frequency distribution of sentiment symbols in the Weibo dataset, which shows that the distribution of sentiment symbols has a long-tail phenomenon, i.e., only a few sentiment symbols appear frequently, while most sentiment symbols are used less and account for the majority of the total. Therefore, sentiment symbols with the number of occurrences greater than the threshold value  $\theta_1 = 200$  are selected to build the sentiment feature model.

The lines in Figure 6 indicate the frequency of emotion symbols in the test data set. The same statistical analysis of the sentiment symbols in the test dataset, as shown in Figure 6, also has a long-tail phenomenon, where most

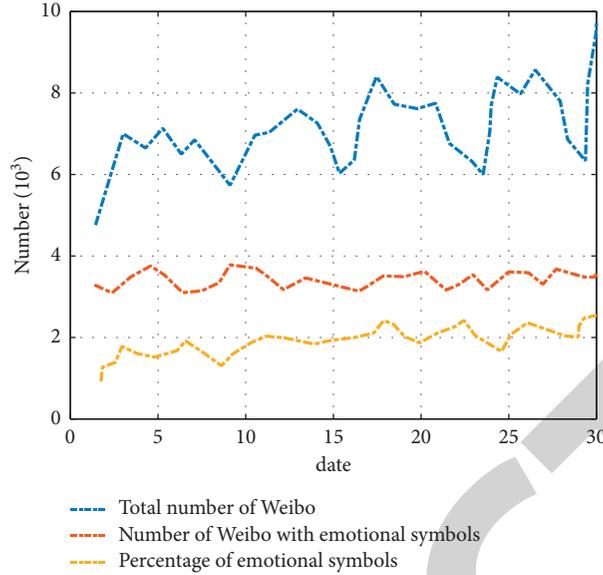


FIGURE 4: Distribution of microblog texts containing emotion symbols.

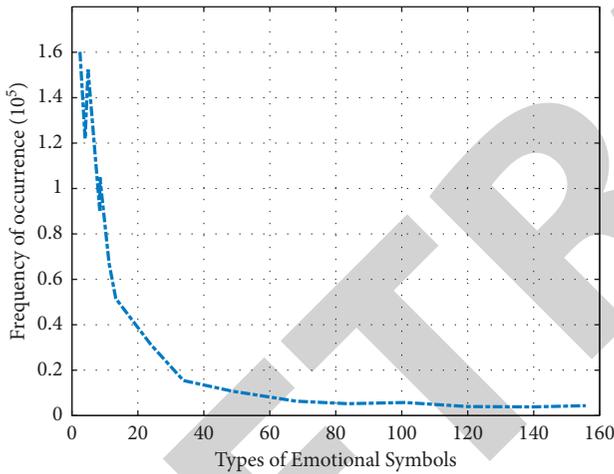


FIGURE 5: Distribution of the frequency of emotional symbols.

sentiment features do not contain microbiology texts in the time window or contain only a very small number of texts, and the sentiments used are only a few. Therefore, mainstream sentiments in the time window are analyzed and extracted for best period detection based on the number of texts in the sentimental text set to reduce the data size and improve the detection efficiency .

**4.2. Experimental Results and Analysis.** In this paper, the time windows are divided by hours, and the data in the corresponding time windows are input into the model. After several experiments, it is verified that the extraction of bursting words works best when the weight factor is  $\alpha = 0.6$ ,  $\beta = 0.2$ ,  $\gamma = 0.2$ .

The best period detected on March 8, 11 : 00-12 : 00 taken as an example, some of these words best weight is shown in

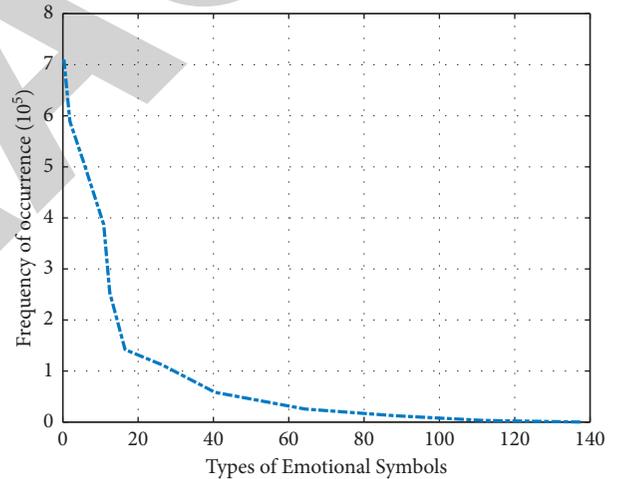


FIGURE 6: Frequency distribution of sentiment symbols in March 2014.

Table 1, which can be seen in the best weight in the table. Figure 7 shows the ten-day word frequency distribution of these words, and it is obvious that an unexpected event occurred on the 8<sup>th</sup>.

In this paper, precision, recall, and  $F1$  values are used as the evaluation metrics of the model. The formula is shown in equations (10) to (12):

$$\text{accuracy} = \frac{\text{number of correctly detected emergencies}}{\text{number of emergencies detected using this model}}, \quad (10)$$

$$\text{recall} = \frac{\text{number of correctly detected emergencies}}{\text{actual number of emergencies in the data set}}, \quad (11)$$

TABLE 1: Partial results of topic label weights.

Words	Word frequency weight	Word frequency growth weight	Topic tag weight	Burst weight of words
Publish	0.8750	1.0000	0.0000	0.72500
Out of contact	1.0000	0.0861	0.1280	0.6492
Flight	1.0000	0.1428	0.00000	0.62856
Malaysia airlines	0.8750	0.2000	0.1480	0.59460
Passenger	0.6250	1.0000	0.0000	0.57500
Name list	0.5000	1.0000	0.0000	0.50000
Safety	0.7500	0.1224	0.0000	0.47448

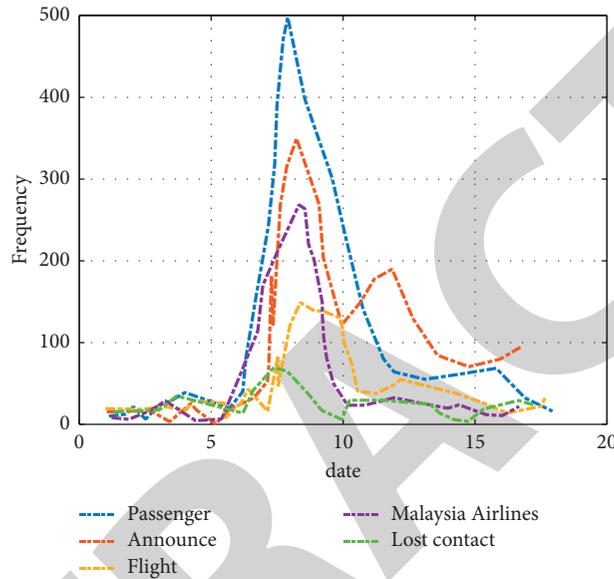


FIGURE 7: Frequency distribution of emergent words.

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}. \quad (12)$$

The intercluster similarity threshold  $\sigma$  is the core parameter for event classification when clustering emergent words. After several experiments, it is verified that the threshold  $\sigma$  takes the value between 13.7 and 14.5 when the clustering effect is better than other values, and the values of this interval are analyzed, and the experimental results are shown in Table 2.

It can be seen that the accuracy, recall, and F1 values reach the highest value when the threshold  $\sigma$  is taken as 14.1, respectively. The reason for this is that as the threshold value decreases, more events are detected, but it is easy to divide the burst words of the same event into multiple clusters resulting in too little information in the clusters to extract the corresponding burst events; while as the threshold value increases, fewer events are detected and the burst words of multiple events are clustered into the same cluster resulting in failure to extract the correct burst events. The intercluster similarity threshold  $\sigma$  is taken as 14.1, and then the best words are clustered to detect the best events, where the results of the correctly detected burst events are partially shown in Table 3.

TABLE 2: Effect of intercluster threshold  $\sigma$  on experimental results.

Threshold $\sigma$	Accuracy	Recall	F1 value
13.7	0.5789	0.7586	0.6567
13.8	0.6111	0.7586	0.6769
13.9	0.7059	0.8276	0.7619
14.0	0.7813	0.8621	0.8197
14.1	0.8361	0.8793	0.8571
14.2	0.8077	0.7541	0.7636
14.3	0.7727	0.5862	0.6661
14.4	0.7222	0.4483	0.5532
14.5	0.6471	0.3793	0.4783

It can be seen that the method proposed in this paper can detect the breaking events in microbiology more accurately. For example, the sentiment feature of “[candle], [tear]” was used to detect the breaking words “railway station”, “Kunming”, “violence”, “death”, “terror” on March 1, 2014. “, “death”, “terror”, and the corresponding outbreak is “Kunming railway station violence”.

Figure 8 shows the total frequency distribution of the sentiment symbols “[candle], [tear]” in March 2014, and three data peaks can be found, corresponding to the “violent incident at Kunming railway station” on the night of the 1st,

TABLE 3: Selected results of emergencies.

Time	Emotional characteristics	Burst word set	Related microblog
2014.3.1	[Candles], [tears]	Railway station, Kunming, violence, death, terror	On the evening of March 1, a mob with knives occurred in Kunming railway station
2014.3.8	[Candles], [tears]	Flight, Malaysia airlines, passengers, announcement, loss of contact	Malaysia airlines lost contact with mh37 flight passenger list announced, onboard
2014.3.13	[Applause]	China, international, government, premier, reporter, Li Keqiang	Premier answered reporters' questions # people's daily, people's daily, people's daily
2014.3.24	Awesome	Champion, Beijing, jinyu, congratulations, Niu Shengli	Congratulations to Beijing jinyu basketball team for winning the CBA League again

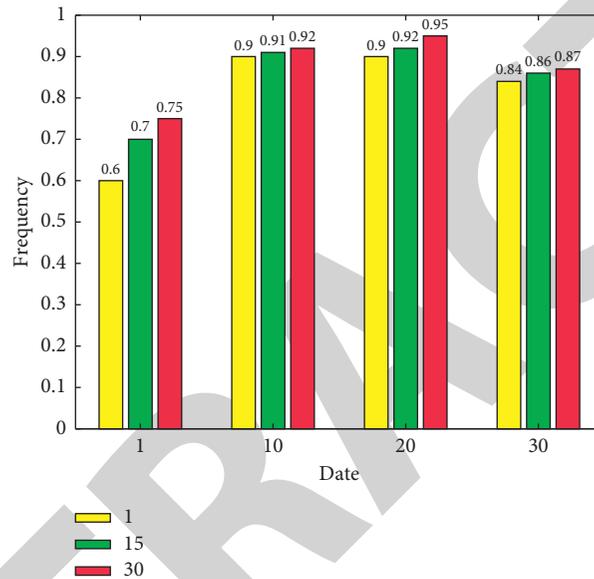


FIGURE 8: Daily distribution of total frequency of “[candle]” and “[tear]” emotion symbols.

the “Malaysia airlines crash” on the 8th, and the “Malaysian Prime Minister announced the crash of Malaysia airlines” on the 24th. “Malaysia airlines crash” on the 1st, “Malaysia airlines crash” on the 8th, and “Malaysian Prime Minister announced the crash of Malaysia airlines” on the 24th. Therefore, when an unexpected event occurs, the corresponding emotional traits will also be in an unexpected state, which also shows the effectiveness of using emotional traits to detect unexpected periods. Specifically, in Figure 8 the frequency of simultaneous occurrence of public events on different dates, for example, the probability of a public event occurring on the 1st of each month is 0.6, which is yellow, and the probability of occurrence of 15 events is 0.7, indicated in green; the probability of occurrence of 30 events is 0.75, indicated in red.

To further verify the effectiveness of the method in this paper, it is compared with the BBW [28] algorithm proposed by Zhang. This algorithm extracts the burst words based on the improved TF-IDF to calculate the base weight and burst weight in each time window, and then clusters to detect burst events based on the obtained burst words,

which is a more classical algorithm; compared with the Burst\_st algorithm proposed by [22], which is a more efficient algorithm among the current feature-centered methods, this algorithm combines sentiment features with microbial topic tags to detect the breaking events, which has some comparability. The experimental comparison results are shown in Table 4.

It can be seen that the method proposed in this paper has a greater improvement in accuracy and F1 value and a smaller improvement in recall. To analyze the reasons, the BBW algorithm needs to calculate the base and burst weights in each time window and then extract the burst words, which cannot fully extract the burst words, while this paper only needs to extract the burst words within the burst period, which can improve the efficiency and identify the burst words more accurately at the same time; the Burst\_st algorithm only detects the burst events based on the topic tags and news headlines in microbiology, while many microbiology originally does not have topic tags, or contain some useless topic tags. In this paper, we introduce the word frequency and word frequency growth rate to improve detection accuracy.

TABLE 4: Comparison of the results of different algorithms.

Algorithm	Accuracy	Recall	F1 value
BBW	0.6341	0.7759	0.6979
Burst_st	0.7551	0.8448	0.7974
MF_BDM	0.8361	0.8793	0.8571

## 5. Conclusions

In this paper, we propose an MCMF-A for microbiology. The results show that the F1 value of the MCMF-A sentiment analysis model reaches 90.21%, which is 9.71% and 9.14% higher than the benchmark CNN and BiLSTM models, respectively. Based on deep learning and multi-channel, this model achieves optimal results by introducing an attention mechanism and fusing the local and global semantic features captured by CNN and BiLSTM, which further promote the progress of microbial sentiment analysis research. The model achieves optimal results by introducing attention mechanisms and fusing local and global semantic features of microbial text captured by CNN and BiLSTM on the basis of deep learning and multichannel, which further advances the research of microbial sentiment analysis.

## Data Availability

The dataset used in this paper is available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

## References

- [1] X. Zhu, B. Fu, Y. Yang et al., "Attention-based recurrent neural network for influenza epidemic prediction," *BMC Bioinformatics*, vol. 20, no. 18, pp. 1–10, 2019.
- [2] Z. Meng, S. Tian, L. Yu, and Y. Lv, "Joint extraction of entities and relations based on character graph convolutional network and multi-head self-attention mechanism," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 33, no. 2, pp. 349–362, 2021.
- [3] D. Shan, W. X. Zhao, R. Chen et al., "Eventsearch: a system for event discovery and retrieval on multi-type historical data," in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1564–1567, Beijing China, August 2012.
- [4] M. Ling, Q. Chen, Q. Sun, and Y. Jia, "Hybrid neural network for Sina Weibo sentiment analysis," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 4, pp. 983–990, 2020.
- [5] K. Chen, Y. Zhou, H. Zha, J. He, P. Shen, and X. Yang, "Cost-effective node monitoring for online hot eventdetection in Sina Weibo microblogging," in *Proceedings of the 22nd International Conference on World Wide Web*, pp. 107–108, Rio de Janeiro, Brazil, May 2013.
- [6] J. Wang, Q. Xu, H. Lin, Z. Yang, and Y. Li, "Semi-supervised method for biomedical event extraction," *Proteome Science*, vol. 11, no. 1, pp. 1–10, 2013.
- [7] H. Tanev, J. Piskorski, and M. Atkinson, "Real-time news event extraction for global crisis monitoring," in *Proceedings of the International Conference on Application of Natural Language to Information Systems*, pp. 207–218, Springer, Saarbrücken, Germany, June 2008.
- [8] S. Xinrui, M. Yu, and W. Wenle, "Identifying traffic events from Weibo with knowledge graph and target detection," *Data Analysis and Knowledge Discovery*, vol. 4, no. 12, pp. 136–147, 2020.
- [9] Y. Yonggan Li, X. Xueguang Zhou, Y. Yan Sun, and Z. Huanguo, "Design and implementation of Weibo sentiment analysis based on LDA and dependency parsing," *China Communications*, vol. 13, no. 11, pp. 91–105, 2016.
- [10] R. Fan, J. Zhao, and K. Xu, "Topic dynamics in Weibo: a comprehensive study," *Social Network Analysis and Mining*, vol. 5, no. 1, pp. 1–15, 2015.
- [11] C. H. Cao, Y. N. Tang, D. Y. Huang, W. Gan, and C. Zhang, "IIBE: An improved identity-based encryption algorithm for wsn security," *Security and Communication Networks*, vol. 2021, Article ID 8527068, 8 pages, 2021.
- [12] Y. Zhang, Y. Chen, C. Huang, and M. Gao, "Object detection network based on feature fusion and attention mechanism," *Future Internet*, vol. 11, no. 1, p. 9, 2019.
- [13] D. Wu, C. Zhang, L. Ji, R. Ran, H. Wu, and Y. Xu, "Forest fire recognition based on feature extraction from multi-view images," *Traitement du Signal*, vol. 38, no. 3, pp. 775–783, 2021.
- [14] M. Ju, J. Luo, Z. Wang, and H. Luo, "Adaptive feature fusion with attention mechanism for multi-scale target detection," *Neural Computing & Applications*, vol. 33, no. 7, pp. 2769–2781, 2021.
- [15] C. Huang, Y. Lan, G. Xu et al., "A deep segmentation network of multi-scale feature fusion based on attention mechanism for IVOCT lumen contour," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 1, pp. 62–69, 2020.
- [16] H. Zhou, Z. Fang, Y. Gao, B. Huang, C. Zhong, and R. Shang, "Feature fusion network based on attention mechanism for 3D semantic segmentation of point clouds," *Pattern Recognition Letters*, vol. 133, pp. 327–333, 2020.
- [17] Y. Cheng, W. Liu, and W. Xing, "Weighted feature fusion and attention mechanism for object detection," *Journal of Electronic Imaging*, vol. 30, no. 2, Article ID 023015, 2021.
- [18] L. Wang, C. Zhang, Q. Chen et al., "A communication strategy of proactive nodes based on loop theorem in wireless sensor networks," in *Proceedings of the 2018 Ninth International Conference on Intelligent Control and Information Processing (ICICIP)*, pp. 160–167, IEEE, Wanzhou, China, November 2018.
- [19] H. Gao, Y. Zhang, Z. Chen, and C. Li, "A multiscale dual-branch feature fusion and attention network for hyperspectral images classification," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 8180–8192, 2021.
- [20] X. Guo, X. Lu, Q. Liu, and X. Che, "EMFN: enhanced multi-feature fusion network for hard exudate detection in fundus images," *IEEE Access*, vol. 7, pp. 176912–176920, 2019.
- [21] Q. Zhou, Z. Qu, and C. Cao, "Mixed pooling and richer attention feature fusion for crack detection," *Pattern Recognition Letters*, vol. 145, pp. 96–102, 2021.
- [22] M. Naughton, N. Kushmerick, and J. Carthy, "Event extraction from heterogeneous news sources," in *Proceedings of the AAAI Workshop Event Extraction and Synthesis*, pp. 1–6, Boston, MA, USA, July 2006.

- [23] V. Erik, "Cluster-centric approach to news event extraction," *New Trends in Multimedia and Network Information Systems*, vol. 181, p. 276, 2008.
- [24] S. Han, X. Hao, and H. Huang, "An event-extraction approach for business analysis from online Chinese news," *Electronic Commerce Research and Applications*, vol. 28, pp. 244–260, 2018.
- [25] H. Li, D. Zeng, L. Chen, Q. Chen, M. Wang, and C. Zhang, "Immune multipath reliable transmission with fault tolerance in wireless sensor networks," in *Proceedings of the International Conference on Bio-Inspired Computing: Theories and Applications*, pp. 513–517, Springer, Xi'an, China, October 2016.
- [26] S. Mutuvi, A. Doucet, G. Lejeune, and M. Odeo, "A dataset for multilingual epidemiological event extraction," in *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, pp. 4139–4144, Marseille, France, June 2020.
- [27] D. Bellos, M. Basham, T. Pridmore, and A. P. French, "Temporal refinement of 3D CNN semantic segmentations on 4D time-series of undersampled tomograms using hidden Markov models," *Scientific Reports*, vol. 11, no. 1, pp. 23279–23314, 2021.
- [28] P. Jin, L. Mu, L. Zheng, and L. Yue, "News feature extraction for events on social network platforms," in *Proceedings of the 26th International Conference on World Wide Web Companion*, pp. 69–78, Perth, Australia, April 2017.
- [29] H. Zhang, F. Boons, and R. Batista-Navarro, "Whose story is it anyway? a," *Information Processing & Management*, vol. 56, no. 5, pp. 1837–1848, 2019.
- [30] Z. Wang, L. Sun, X. Li, and L. Wang, "Event extraction via dmcnn in open domain public sentiment information," in *Proceedings of the International Conference of Pioneering Computer Scientists, Engineers and Educators*, pp. 90–100, Springer, Taiyuan, China, September 2020.