Front. Comput. Sci., 2023, 17(2): 172314
https://doi.org/10.1007/s11704-022-1409-x

RESEARCH ARTICLE

Pairwise tagging framework for end-to-end emotion-cause pair extraction

Zhen WU[1,2], Xinyu DAI (✉)[1,2], Rui XIA[3]

1 National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China
2 Collaborative Innovation Center of Novel Software Technology and Industrialization, Nanjing 210023, China
3 School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210023, China

Abstract Emotion-cause pair extraction (ECPE) aims to extract all the pairs of emotions and corresponding causes in a document. It generally contains three subtasks, emotions extraction, causes extraction, and causal relations detection between emotions and causes. Existing works adopt pipelined approaches or multi-task learning to address the ECPE task. However, the pipelined approaches easily suffer from error propagation in real-world scenarios. Typical multi-task learning cannot optimize all tasks globally and may lead to suboptimal extraction results. To address these issues, we propose a novel framework, Pairwise Tagging Framework (PTF), tackling the complete emotion-cause pair extraction in one unified tagging task. Unlike prior works, PTF innovatively transforms all subtasks of ECPE, i.e., emotions extraction, causes extraction, and causal relations detection between emotions and causes, into one unified clause-pair tagging task. Through this unified tagging task, we can optimize the ECPE task globally and extract more accurate emotion-cause pairs. To validate the feasibility and effectiveness of PTF, we design an end-to-end PTF-based neural network and conduct experiments on the ECPE benchmark dataset. The experimental results show that our method outperforms pipelined approaches significantly and typical multi-task learning approaches.

Keywords emotion-cause pair extraction, pairwise tagging framework, end-to-end, neural network

1 Introduction

Text emotion analysis aims to analyze the emotion of people from text and plays a crucial role in human communication and making decisions [1,2]. The early studies mainly focus on emotion classification that aims to infer emotion category of user-generated text [3–5]. Beyond emotion classification, people sometimes care more about the causes stimulating a specific emotion. Therefore, a fine-grained emotion analysis task, emotion cause extraction (ECE), is proposed to extract the corresponding causes that lead to a given emotion expression in text [6]. In recent years, the ECE task has drawn increasing attention due to its potential applications [7–12]. However, this popular task requires that emotions must be annotated in advance before cause extraction, which limits its applications in real-world scenarios. To break the restriction, a new emotion analysis task, emotion-cause pair extraction (ECPE), is proposed to extract all the pairs of emotions and corresponding causes in a document [13]. Different from ECE, ECPE can extract emotions-cause pairs automatically without additional emotion annotation.

The ECPE is a fine-grained emotion analysis task, and Xia et al. [13] formalize it as a clause pair extraction problem. In this task, a clause expressing a certain emotion is regarded as an emotion clause, and the clauses stimulating this emotion are corresponding cause clauses. Figure 1 shows an example of the ECPE task. There are six clauses in the example document. The fourth clause and the fifth clause are two *emotion clauses* because they respectively express the emotion "happy" and "worried". The emotion "happy" has two corresponding causes, respectively "a policeman visited the old man with the lost money" in the second clause and "told him that the thief was caught" in the third clause. The cause of the emotion "worried" is the sixth clause "as he doesn't know

Clause c_1: Yesterday morning,
Clause c_2: a policeman visited the old man with the lost money,
Clause c_3: and told him that the thief was caught.
Clause c_4: **The old man was very happy**.
Clause c_5: **But he still feels worried**,
Clause c_6: as he doesn't know how to keep so much money.

Emotion-cause pair 1: (**The old man was very happy**, a policeman visited the old man with the lost money)

Emotion-cause pair 2: (**The old man was very happy**, and told him that the thief was caught)

Emotion-cause pair 3: (**But he still feels worried**, as he doesn't know how to keep so much money)

Fig. 1 An example of the ECPE task. There are six clauses in the document. The clauses highlighted in bold are emotion clauses, and the clauses in underline are cause clauses. The goal of ECPE is to extract three emotion-cause pairs shown in the lower part

Received July 28, 2021; accepted December 2, 2021

E-mail: daixinyu@nju.edu.cn

how to keep so much money". Therefore, the goal of ECPE is to directly extract the three emotion-cause pairs from the document, i.e., ("The old man was very happy", "a policeman visited the old man with the lost money"), ("The old man was very happy", "and told him that the thief was caught"), and ("But he still feels worried", "as he doesn't know how to keep so much money").

To address this new task, Xia et al. [13] propose a two-step solution, which first extracts a set of emotion clauses and a set of cause clauses via multi-task learning, then applies Cartesian product to generate all candidate pairs and train a binary filter to eliminate the pairs without a causal relation. Despite the effectiveness of the pipelined approach, it unavoidably suffers from error propagation. Specifically, the extraction errors of emotion or cause in the first step will directly harm the pairing results of the second step. Thereafter, more studies adopt multi-task learning to extract emotion-cause pairs in the end-to-end fashion [14–17]. These multi-task approaches benefit from auxiliary tasks of emotion extraction and cause extraction and achieve performance improvement. Nevertheless, the multiple subtasks in multi-task approaches are optimized in a joint instead of a global way, which means they may not achieve the best potential simultaneously and lead to suboptimal extraction results of emotion-cause pairs. Therefore, one more promising solution is to solve all subtasks of ECPE with a global and unified task. However, this goal is very challenging because the emotion/cause extraction and causal relation detection between emotion and cause are two subtasks of different types and granularity in ECPE. Specifically, the former is a clause-level extraction problem, while the latter is a clause-pair classification problem, which makes it hard to integrate them into one unified task.

To achieve the above challenging goal, we propose a novel framework, Pairwise Tagging Framework (PTF), successfully extracting emotion-cause pairs with one global and unified tagging task. In this novel framework, we innovatively transform emotion extraction and cause extraction into relations detection of clause-pairs, instead of classical sequence extraction [14,16]. In this way, all subtasks of ECPE, including emotion extraction, cause extraction, and causal relations between emotions and causes can be integrated into one unified clause-pair tagging task. Naturally, they can be extracted simultaneously and optimized globally. To validate the feasibility of the proposed framework, we specially design an end-to-end neural network based on PTF. Empirically, we conduct experiments on the ECPE benchmark dataset and the results demonstrate the effectiveness of PTF.

The main contributions of this work are as follows:

- We propose a novel tagging framework PTF for the ECPE task. To the best of our knowledge, PTF is the first work to solve all subtasks of ECPE with a global and unified clause-pair tagging task, instead of multi-task learning.
- We develop a PTF-based end-to-end neural network to validate the effectiveness of PTF. In this neural model, we introduce some helpful mechanisms to improve the performance of ECPE further.
- The experimental results show that our method achieves the obvious improvements compared to previous works and achieves the state-of-the-art performance on the ECPE benchmark dataset.

2 Related work

2.1 Emotion cause extraction

Prior works mainly focus on emotion cause extraction (ECE), which aims to extract the corresponding causes leading to a given emotion from text. The ECE task was first proposed by [6] and defined as a word-level extraction problem. Early research devotes to designing rule-based methods [8,9,18–21], or using traditional machine learning algorithms such as SVM [22,23] and CRF [23,24] to address this task. Chen et al. [22] realized the difficulty of describing emotion causes at the word level, thus they suggested that a clause may be the more appropriate unit to describe causes. Following the idea, Gui et al. [10] built a clause-level ECE dataset using SINA city news. In this corpus, an annotation of emotion clause corresponds to one or multiple cause clauses. Recently, it has received much attention and become a benchmark dataset for ECE research. Based on the dataset, Gui et al. [25] and Xu et al. [26] respectively proposed structure representation describing the events and multi-kernel learning to extract emotion cause clauses. In the following studies, deep learning techniques have also been applied to the ECE task, such as long short-term memory [11], deep memory network [27], co-attention neural network [28], and three-level (word-phrase-clause) hierarchical network [29]. However, these works ignored the importance of relative position and global label information for emotion cause identification. To address the two issues, Ding et al. [30] converted the task to a reordered clause classification problem. Xia et al. [31] proposed RNN-Transformer based framework to integrate relative position and global prediction information.

2.2 Emotion-cause pair extraction

The ECE task has been studied for about a decade, while it needs additional emotion annotation, which limits its applications in real-world scenarios. To break the limitation, Xia et al. [13] proposed a new emotion analysis task, emotion-cause pair extraction (ECPE), to extract all pairs of emotions and corresponding causes in a document without emotion annotation in advance. They proposed a two-step approach to perform ECPE, which easily suffered from error propagation. To address the issue, more works adopt end-to-end multi-task learning to solve the ECPE task. For example, Song et al. [14] regarded pair extraction as a link prediction task using a vanilla multi-task framework. Wu et al. [32] employs shared encoder and private encoder to improve the performance of subtasks of emotion extraction and cause extraction. Tang et al. [15] proposed a multi-level attention mechanism to capture the word-level and clause-level dependency relations for extracting emotion-cause pairs. Ding et al. [16] employed a 2D Transformer to model the interactions of different emotion-cause pairs and integrated representation, interaction, and prediction into a multi-task framework. On the basis of multi-task learning, Fan et al. [33] further introduced tag distribution refinement strategy to benefit emotion-cause pair extraction using the output of the two auxiliary tasks. Cheng et al. [34] introduced local pair searcher (LPS) to find the corresponding emotions/causes from the nearby clauses of a cause or emotion. Yu et al. [35] designed an additional self-distillation strategy to boost emotion-cause pair extraction. Wei et al. [17] tackled emotion-cause pair extraction from a ranking perspective and employed an external sentiment lexicon to help recall emotion-cause pairs. Besides multi-task learning, there are also some other strategies proposed for the ECPE task. Fan et al. [36] adopted a transition-based strategy and transformed ECPE into a task of parsing-like directed graph construction. Yuan et al. [37] and Cheng et al. [38] constructed different label sets and used sequence labeling scheme to extract emotion-cause pair.

3 Pairwise tagging framework

In this section, we first give the task definition of emotion-cause pairs extraction (ECPE), then explain how the ECPE is represented in Pairwise Tagging Framework (PTF). Finally, we present how to decode emotion-cause pairs according to the tagging results of PTF.

3.1 Task definition

Formally, given a document d consisting of n clauses, i.e., $d = \{c_1, c_2, \ldots, c_n\}$, where c_i is the ith clause of the document d, the goal of the ECPE task is to extract a set P of emotion-cause pairs from the document d:

$$P = \{\ldots, (c_e, c_c), \ldots\}, \tag{1}$$

where (c_e, c_c) is an emotion-cause pair, the notation c_e represents an emotion clause and c_c denotes the corresponding cause clause of c_e.

3.2 Pairwise tagging

As aforementioned, we transform emotion extraction, cause extraction, and causal relation detection between any two clauses into a clause-pairs tagging task. Thus the whole ECPE task can be solved with one unified tagging task.

Specifically, we use four tags $\{\mathtt{E}, \mathtt{C}, \mathtt{P}, \mathtt{O}\}$ to denote the relations of any one unordered clause-pair (c_i, c_j) $(1 \leqslant i, j \leqslant n)$ in a document d. Their meanings are shown in Table 1. The tags E and C are applied in the position of the main diagonal. The tag E is used to represent an emotion expressed by the clause pair (c_i, c_i), i.e., the clause c_i is an emotion clause. Correspondingly, the clause pair (c_j, c_j) with the tag C indicates the clause c_j is a cause clause. The tag P represents that the clause-pair (c_i, c_j) contains a causal relation. Note that, the tag P cannot indicate which clause is an emotion clause or a cause clause in the clause pair (c_i, c_j), because the relation of the pair (c_i, c_j) is unordered in PTF. The tag O denotes no any above relations between two clauses.

Table 1 The meanings of PTF tags for the ECPE task

Tags	Meanings
E	the clause-pair (c_i, c_i) expresses emotion, i.e., c_i is an emotion clause.
C	the clause-pair (c_j, c_j) expresses cause, i.e., c_j is a cause clause.
P	the clause-pair (c_i, c_j) contains a casual relation.
O	no above three relations for clause-pair (c_i, c_j).

We take the document of Fig. 1 for example to elaborate our PTF tagging. As shown in Fig. 2, there are two emotion clauses c_4 and c_5, three cause clauses c_2, c_3, and c_6, three causal clause pairs (c_2, c_4), (c_3, c_4) and (c_5, c_6) in the document. Therefore, we tag (c_4, c_4) and (c_5, c_5) with E, (c_2, c_2), (c_3, c_3), and (c_6, c_6) with C, (c_2, c_4), (c_3, c_4), and (c_5, c_6) with P, and other clause-pairs with the tag O.

3.3 Decoding algorithm

Given a document d, we can obtain the predicted PTF tagging results T using traditional machine learning algorithms or neural networks. We adopt a neural network model in this work (See Section 4). In this subsection, we focus on how to decode the final emotion-cause pairs from the predicted tagging results T. $T(c_i, c_j)$ represents the predicted tag of the clause-pair (c_i, c_j). Algorithm 1 shows the decoding details.

As we have mentioned, the relation of clause pair (c_i, c_j) in PTF is unordered, so the clause pair (c_i, c_j) with the predicted tag P cannot be directly used to extract emotion-cause pairs. To address this issue, we first obtain the set E of emotion

	c_1	c_2	c_3	c_4	c_5	c_6	
	O	O	O	O	O	O	c_1
		C	O	P	O	O	c_2
			C	P	O	O	c_3
				E	O	O	c_4
					E	P	c_5
						C	c_6

Fig. 2 PTF tagging result of example document of Fig. 1. There are three emotion-cause pairs (c_4, c_2), (c_4, c_3), and (c_5, c_6) in the example document

Algorithm 1 Decoding algorithm

Input: The tagging reuslts T in PTF. $T(c_i, c_j)$ represents the predicted tag of the clause-pair (c_i, c_j).

Output: Emotion-cause pair set P

```
Let E ← ∅, C ← ∅, P ← ∅.
while a clause c_i in d do
    if T(c_i, c_i) = E then
        E ← E ∪ {c_i}.
    end if
end while
while a clause c_j in d do
    if T(c_j, c_j) = C then
        C ← C ∪ {c_j}.
    end if
end while
while c_e in E and c_c in C do
    if T(c_e, c_c) = P then
        P ← P ∪ {(c_e, c_c)}.
    end if
end while
return the set P
```

clauses and the set C of cause clauses from the tagging result T. As shown in line 2 to line 11 of Algorithm 1, we regard the clause c_i as an emotion clause when the $T(c_i,c_i)$ is predicted as the tag E. Similarly, $T(c_j,c_j)$ predicted as the tag C represents that the clause c_j is a cause clause. Afterwards, we judge an emotion clause c_e of the set E and a cause clause c_c of the set C as an emotion-cause pair if $T(c_e,c_c)$ is predicted as the tag P, as shown in line 12 to line 16 of Algorithm 1.

Through the above pairwise clause-pair tagging and the decoding algorithm, the complete ECPE task can be solved with our unified PTF tagging task, instead of several subtasks.

4 Model

In this section, we design an end-to-end neural model Pairwise Tagging Network (PTN), as the validation system, to demonstrate the feasibility and effectiveness of PTF for the ECPE task. Figure 3 shows an illustration of PTN. We will first describe its clause-pair representations learning module. Then we introduce three information-enhancing mechanisms to enhance the clause-pair representations. Finally, we present the decoding and training modules of PTN.

4.1 Clause pair representations learning

Given a document containing n clauses $d=\{c_1,c_2,\dots,c_n\}$, its i-th clause c_i contains l_i words $c_i=\{w_{i1},w_{i2},\dots,w_{il_i}\}$, we first map each word w_{it} into a word vector $\mathbf{w}_{it}$ by looking up an embedding table $\mathbf{E}_{emb}\in\mathbb{R}^{m\times|V|}$, where m is the embedding dimension and $|V|$ denotes the vocabulary size.

Following the hierarchical document modeling works [39,40], we then adopt a word-level bidirectional Long Short-Term Memory (LSTM) [41] to model the internal semantics of each clause. For the ith clause c_i, it encodes the word vector sequence $\{\mathbf{w}_{i1},\mathbf{w}_{i2},\dots,\mathbf{w}_{il_i}\}$ and generate the corresponding hidden states $\{\mathbf{h}_{i1},\mathbf{h}_{i2},\dots,\mathbf{h}_{il_i}\}$. To obtain more effective clause representation, we introduce the attention mechanism [42] to aggregate the informative words instead of using pooling or the last hidden state. The representation $\mathbf{c}_i$ of the clause c_i can be obtained as follows:

$$g(\mathbf{h}_{it})=(\mathbf{v}_w)^\top\tanh(\mathbf{W}_w\mathbf{h}_{it}+\mathbf{b}_w), \tag{2}$$

$$\alpha_{it}=\frac{\exp(g(\mathbf{h}_{it}))}{\sum_{k=1}^{n}(g(\mathbf{h}_{ik}))}, \tag{3}$$

$$\mathbf{c}_i=\sum_{t=1}^{l_i}\alpha_{it}\mathbf{h}_{it}, \tag{4}$$

where $\mathbf{W}_w$ is the weight matrix, $\mathbf{v}_w$ indicates the weight vector, and $\mathbf{b}_w$ denotes the bias.

We can observe that the above representations of different clauses are mutually independent, which is unfavorable for modeling relations between clauses. To capture their dependency, we employ a clause-level BiLSTM network to encode the clause representations $\{\mathbf{c}_1,\mathbf{c}_2,\dots,\mathbf{c}_n\}$ and generate clause-level hidden states $\{\mathbf{h}_1,\mathbf{h}_2,\dots,\mathbf{h}_n\}$.

Finally, we concatenate $\mathbf{h}_i$ and $\mathbf{h}_j$ as the representation of the clause pair (c_i,c_j), i.e., $\mathbf{r}_{ij}=[\mathbf{h}_i;\mathbf{h}_j]$, where $[\cdot;\cdot]$ denotes the vector concatenation operation, to detect relations between the clauses c_i and c_j.

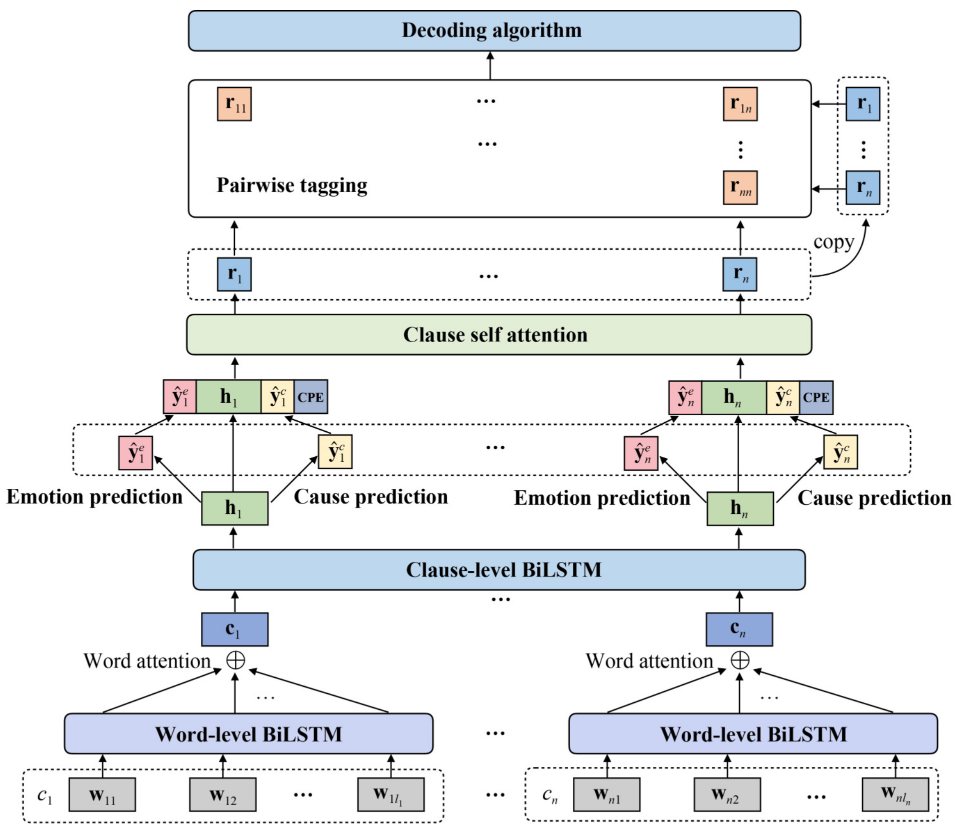

Fig. 3 An illustration of pairwise tagging network (PTN)

4.2 Information-enhancing mechanisms

To better capture the relations between clauses, we additionally design three information-enhancing mechanisms, namely clause position embedding (CPE) and emotion/cause prediction (ECP) and clause self attention (CSA), to enhance the clause pair representation $\mathbf{r}_{ij}$.

4.2.1 Clause position embedding

Firstly, some works [30,31] reveal that a clause closer to the emotion is more likely to be emotion's cause. In other words, two closer clauses in a document are more likely to be an emotion-cause pair, otherwise less likely. To leverage this important indicative information, we introduce clause position embedding to represent them. Specifically, the position indicators $\{1,2,\ldots,n\}$ of clauses $\{c_1,c_2,\ldots,c_n\}$ are mapped into continuous representations $\{\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_n\}$. Then we concatenate clause position embedding $\mathbf{p}_i$ and the clause representation $\mathbf{h}_i$ as the enhanced clause representation:

$$\widetilde{\mathbf{h}}_i = [\mathbf{h}_i;\mathbf{p}_i]. \tag{5}$$

4.2.2 Emotion/cause prediction

Secondly, an emotion-cause pair contains an emotion clause and a cause clause. Correspondingly, if a clause is an emotion clause or a cause clause, it is quite likely to be part of an emotion-cause pair. Therefore, if we predict the emotion probability and the cause probability of a clause in advance, these predictions ought to be helpful for detecting the relation of clause pair. To this end, we use the original clause representation $\mathbf{h}_i$ to predict emotion probability $\hat{\mathbf{y}}_i^e$ and cause probability $\hat{\mathbf{y}}_i^c$ of the clause c_i. Their prediction probabilities and the corresponding losses are defined as follows:

$$\hat{\mathbf{y}}_i^e = \mathrm{softmax}(\mathbf{W}_\mathrm{e}\mathbf{h}_\mathrm{i} + \mathbf{b}_\mathrm{e}), \tag{6}$$

$$loss_e = -\sum_{i=1}^{n}\sum_{k=0}^{1} y_{ik}^e \cdot \log(\hat{y}_{ik}^e), \tag{7}$$

$$\hat{\mathbf{y}}_i^c = \mathrm{softmax}(\mathbf{W}_\mathrm{c}\mathbf{h}_\mathrm{i} + \mathbf{b}_\mathrm{c}), \tag{8}$$

$$loss_c = -\sum_{i=1}^{n}\sum_{k=0}^{1} y_{ik}^c \cdot \log(\hat{y}_{ik}^c), \tag{9}$$

where $\mathbf{y}_i^e$ and $\mathbf{y}_i^c$ respectively denote the ground truth distribution of emotion and cause.

To utilize the above helpful emotion/cause predictions, we update the enhanced clause representation $\widetilde{\mathbf{h}}_i$ as follows:

$$\widetilde{\mathbf{h}}_i = [\mathbf{h}_i;\mathbf{p}_i;\hat{\mathbf{y}}_i^e;\hat{\mathbf{y}}_i^c]. \tag{10}$$

4.2.3 Clause self attention

Thirdly, it is important for the ECPE task to capture the potential connection of clause-pairs. As we all know, self attention is a good mechanism to caputer the internal connection of a sequence [43]. Therefore, we use a clause self attention mechanism on the representations $\{\widetilde{\mathbf{h}}_1,\widetilde{\mathbf{h}}_2,\ldots,\widetilde{\mathbf{h}}_n\}$ to model the relevance between clauses. The importance β_{ij} of the jth clause c_j for the ith clause c_i can be defined as:

$$g(\widetilde{\mathbf{h}}_i,\widetilde{\mathbf{h}}_j) = (\mathbf{v}_s)^\top \tanh(\mathbf{W}_{s_1}\widetilde{\mathbf{h}}_i + \mathbf{W}_{s_2}\widetilde{\mathbf{h}}_j + \mathbf{b}_s), \tag{11}$$

$$\beta_{ij} = \frac{\exp(g(\widetilde{\mathbf{h}}_i,\widetilde{\mathbf{h}}_j))}{\sum_{k=1}^{n}(g(\widetilde{\mathbf{h}}_i,\widetilde{\mathbf{h}}_k))}, \tag{12}$$

where $\mathbf{W}_{s_1}$ and $\mathbf{W}_{s_2}$ are weight matrices, $\mathbf{v}_s$ is the weight vector, and $\mathbf{b}_s$ represents the bias.

The meaning of attention weight β_{ij} is different from that of α_{it} in Eq. (3). The α_{it} indicates the importance of the word w_{it} for the whole semantics of the clause c_i, which aims to aggregate information to obtain clause representation. In contrast, the β_{ij} is used to measure the relevance of the clauses c_i and c_j. We then exploit β_{ij} to obtain the related clause information of the ith clause and inject them into the representation of c_i as follows:

$$\mathbf{r}_i = \widetilde{\mathbf{h}}_i + \sum_{j=1}^{n}\beta_{ij}\widetilde{\mathbf{h}}_j. \tag{13}$$

Finally, we replace pair representation $\mathbf{r}_{ij} = [\mathbf{h}_i;\mathbf{h}_j]$ with information-enhanced representation $\mathbf{r}_{ij} = [\mathbf{r}_i;\mathbf{r}_j]$ for detecting the relation of the clause-pair (c_i,c_j).

4.3 Decoding and training

To extract emotion-cause pairs, we first employ a linear layer and a softmax layer to predict the relation distribution of the clause-pair (c_i,c_j):

$$\hat{\mathbf{y}}_{ij} = \mathrm{softmax}(\mathbf{W}_p\mathbf{r}_{ij} + \mathbf{b}_p), \tag{14}$$

where $\mathbf{W}_p$ and $\mathbf{b}_p$ respectively represent the weight matrix and bias. The cross-entropy loss of PTF tagging on all clause-pairs is defined as $loss$:

$$loss = -\sum_{i=1}^{n}\sum_{j=i}^{n}\sum_{k=0}^{3} y_{ijk}\cdot\log(\hat{y}_{ijk}), \tag{15}$$

where $\mathbf{y}_{ij}$ is the ground truth tag of (c_i,c_j), the tags $\{\mathtt{O},\mathtt{E},\mathtt{C},\mathtt{P}\}$ are correspondingly mapped into the labels $\{0,1,2,3\}$. We can obtain the tagging results T from the predicted probability distribution $\hat{\mathbf{y}}$ by $\mathtt{argmax}(\hat{\mathbf{y}})$ and then decode the all emotion-cause pairs according to the Algorithm 1.

To ensure the predicted distribution $\hat{\mathbf{y}}_i^e$ in Eq. (6) and $\hat{\mathbf{y}}_i^c$ in Eq. (8) contain the indications of emotion and cause, we additionally optimize the losses $loss_e$ and $loss_c$. Therefore, the final training objective of the model is to minimize the following loss:

$$\mathcal{L} = loss + \lambda_e loss_e + \lambda_c loss_c, \tag{16}$$

where λ_e and λ_c are trade-off factors.

5 Experiments

5.1 Dataset and metrics

We evaluate different methods on the widely used ECPE benchmark dataset [13], which is constructed from the ECE benchmark corpus [25]. Table 2 shows the statistics of the dataset.

Following previous works [13–17], we stochastically select 90% of the data for training and the remaining 10% for testing. To obtain statistically credible results, we repeat each experiment 20 times and report the average result. We use the

Table 2 Statistics of the ECPE dataset

	Number
Documents	1945
Emotions	2085
Causes	2142
Emotion-cause pairs	2167
Doc. with one emotion-cause pair	1746
Doc. with two emotion-cause pairs	177
Doc. with more than two emotion-cause pairs	2

precision (P), recall (R), and F1-score (F1) as the metrics for evaluating different methods. The higher P, R, and F1-score represent the better performance. Note that, F1-score is our main metric because it balances the precision and recall of the predicted results. A predicted emotion-cause pair is deemed correct only if its emotion and the corresponding cause are both predicted correctly. When extracting emotion-cause pairs with the unified PTF tagging task, we can also obtain emotions and causes of the document simultaneously. Therefore, we also evaluate the performance of different methods on two subtasks, i.e., emotion extraction and cause extraction, using P, R, and F1 as the metrics.

5.2 Experimental settings

Table 3 shows the detailed hyper-parameter settings. Specifically, we initialize word vectors with 200-dimension word embeddings that are pre-trained on 1.1 million Weibo corpora [13] with the word2vec toolkit [44]. The word vectors are fixed and not fine-tuned during the training stage. The dimensions of LSTM cell and clause position embedding are respectively set to 100 and 50. The dropout [45] is applied to word embeddings layer with the probability of 0.5. We employ the Adam optimizer [46] to train neural models with a mini-batch size of 32 and a learning rate of 0.005. The maximum numbers of words in each clause and clauses in each document are respectively set to be 75 and 45. The loss weights λ_e and λ_c in Eq. (16) are both set to 1.

5.3 Compared methods

Previous works adopt pipeline or end-to-end multi-task approaches to address the ECPE task. In this work, we compare them with our proposed PTN model.

Xia et al. [13] proposed three different pipelined methods based on the two-step strategy. In the first step, they first extract an emotion set and a cause set from a document via multi-task learning. In the second step, they apply a Cartesian product to yield all candidate emotion-cause pairs and then train a filter to eliminate the pairs without a causal relation. These three methods have the same operation in the second step. The differences in the first step are as follows:

Table 3 Hyper-parameter settings

Hyper-parameters	Values
Dimension of word embedding	200
Dimension of LSTM cell	100
Dimension of position embedding	50
Dropout rate	0.5
Batch size	32
Learning rate	0.005
Maximum words of a clause	75
Maximum clauses of a document	45
Loss weight λ_e	1
Loss weight λ_c	1

- **Indep**: Indep first adopts a BiLSTM to encode word vectors and obtain clause representations, then employs two independent clause BiLSTMs to model contextual information between clauses, respectively for emotion extraction and cause extraction.
- **Inter-CE**: Different from Indep, Inter-CE uses predictions of cause extraction to improve emotion extraction.
- **Inter-EC**: In contrast, Inter-EC uses predictions of emotion extraction to enhance cause extraction.

The end-to-end approaches are as below:

- **E2EECPE** [14]: E2EECPE is a typical multi-task learning method for the ECPE task. It regards emotion-cause pair extraction as a link prediction task and uses auxiliary emotion extraction and cause extraction to boost the pair extraction.
- **ECPE-2D** [16]: Based on multi-task learning, ECPE-2D employs multi-layers transformer [47] to further capture the connection between clauses.
- **TransECPE** [36]: Different from multi-task methods, TransECPE is a transition-based model and transforms the ECPE task into a procedure of parsing-like directed graph construction.
- **RankCP** [17]: RankCP is a state-of-the-art method and tackles ECPE from a ranking perspective and emphasizes inter-clause modeling. It directly regards the predicted top-1 pair as the emotion-cause pair. Besides, it additionally employs a post-processing strategy. For the candidate top-(N−1) pairs, if the pair contains sentiment words according to an external sentiment lexicon, RankCp extracts it as the emotion-cause pair.

5.4 Evaluation and comparison

Table 4 shows the main experimental results of different methods on the ECPE task and two subtasks of emotion extraction and cause extraction.

Among all the pipelined methods, Indep achieves the worst performance on the ECPE task, because it ignores the mutual indication between emotions and causes. With the help of cause predictions, Inter-CE gets better results than Indep on the emotion extraction, thereby boosting emotion-cause pair extraction. On the contrary, Inter-EC benefits from the predictions of emotion extraction and outperforms Indep and Inter-CE significantly on the cause extraction. The significant improvement of cause extraction makes Inter-EC obtain the best ECPE performance in the three pipelined methods. However, Inter-CE and Inter-EC cannot achieve simultaneous improvements on the two subtasks due to sequential dependency. The inconsistent improvements and error propagation of the pipeline strategy finally limit the overall performance of the ECPE task.

In contrast, the end-to-end methods effectively reduce error propagation and outperform the pipelined methods significantly on emotion-cause pair extraction. In these methods,

Table 4 Main experiment results of different methods (%). Best and second-best results are respectively in bold and underline

Models	Emotion extraction			Cause extraction			Emotion-cause pair extraction		
	P	R	F1	P	R	F1	P	R	F1
Indep	83.75	80.71	82.10	69.02	56.73	62.05	68.32	50.82	58.18
Inter-CE	84.94	81.22	83.00	68.09	56.34	61.51	69.02	51.35	59.01
Inter-EC	83.64	81.07	82.30	70.41	60.83	65.07	67.21	57.05	61.28
E2EECPE	85.95	79.15	82.38	70.62	60.30	65.03	64.78	61.05	62.80
ECPE-2D(base)	85.37	81.97	83.54	71.51	62.74	66.76	71.73	57.54	63.66
ECPE-2D	85.12	82.20	83.58	**72.72**	62.98	67.38	69.60	61.18	64.96
TransECPE	80.80	**84.39**	82.56	67.42	65.34	66.36	65.15	63.54	64.34
RankCP(top-1)	87.35	81.46	84.28	71.30	64.68	67.90	69.10	62.54	65.62
RankCP	**87.03**	84.06	**85.48**	69.27	**67.43**	**68.24**	66.98	**65.46**	66.10
PTN(base)	84.12	81.61	82.82	72.02	63.66	67.50	71.38	59.48	64.80
PTN	84.47	82.78	83.60	71.75	64.70	67.99	**76.00**	59.18	**66.50**

E2EECPE and ECPE-2D(base) are two vanilla multi-task models, and the difference is that ECPE-2D(base) exploits auxiliary predictions of emotion and cause to help detect emotion-cause pairs. Based on ECPE-2D(base), ECPE-2D adopts an additional multi-layer transformer to capture the relevance between clauses, thus achieving further performance improvements. The transition-based system TransECPE can capture interdependence between emotions and causes more effectively for pairs prediction and thus shows competitive results compared with multi-task learning models. RankCP(top-1) models inter-clause relationships to learn clause representations using stacked graph attention layers, and integrates relative position enhanced clause pair ranking into the neural network to extract emotion-cause pairs. On the basis of RankCP(top-1), RankCP uses a post-processing strategy and employs an external sentiment lexicon to effectively improve the recall of emotion-cause pair extraction, thus achieving state-of-the-art performance.

PTN(base) is our vanilla implementation and only utilizes the proposed unified tagging scheme PTF to address the ECPE task, without using the information-enhancing mechanisms mentioned in Section 4.2. We can observe that the vanilla PTN(base) outperforms the typical multi-task learning models, such as E2EECPE and ECPE-2D(base). This superiority indicates that our unified tagging scheme PTF optimizes the extractions of each element in the ECPE task from the global perspective, thus obtaining better results of emotion-cause pair extraction. With the help of the information-enhancing strategies, PTN obtains further improvements and achieves the state-of-the-art performance on the emotion-cause pair extraction, which validates the effectiveness of the information-enhancing mechanisms. We can see that PTN still outperforms the RankCP model by 0.4% in F1-score for the ECPE task, even though it does not use post-processing strategy and external sentiment lexicon. Besides, we find that PTN achieves better precision than recall on the emotion-cause pairs extraction. It is reasonable because there are 89.76% of documents contain only one emotion-cause pair in the ECPE dataset. The pair sparsity implies the extremely unbalanced label distribution in our PTF tagging scheme, which makes the model tend to decode fewer pairs and leads to a lower recall. Therefore, some label-balanced training algorithms may help our method achieve better performance. Among the compared methods, some works implement their models based on not only BiLSTM but also the pre-trained BERT [43]. For a fair comparison, we also implement PTN using the pre-trained BERT as encoder, and Table 5 shows the corresponding results. We can observe that PTN(BERT) outperforms other methods significantly on the task of emotion-cause pair extraction and achieves the state-of-the-art performance among the BERT-based methods.

5.5 Ablation study

To investigate the effect of the information-enhancing mechanisms clause position embedding (CPE), emotion/cause prediction probability (ECP), and clause self attention (CSA) on the model PTN, we respectively remove them from PTN to conduct experiments. The results are shown in Table 6.

Compared to the complete model PTN, the ablation models PTN w/o CPE and PTN w/o ECP perform worse on the ECPE task. Their F1-scores drop by 0.67% and 0.81% respectively under 20 repeated experiments, which indicates that clause position embedding and emotion/cause prediction probability are both beneficial for modeling causal relation between clauses. Besides, we observe, among three information-enhancing mechanisms, clause self attention is the most effective for improving the ECPE task, as the performance of PTN w/o CSA declines by about 1.32% in F1-score after removing clause self attention mechanism from PTN.

5.6 Case study

To show the advantage of our PTN compared with the state-

Table 5 Experiment results of different methods using BERT as the encoder (%). Best and second-best results are respectively in bold and underline

Models	Emotion extraction			Cause extraction			Emotion-cause pair extraction		
	P	R	F1	P	R	F1	P	R	F1
ECPE-2D(BERT)	86.27	**92.21**	89.10	73.36	69.34	71.23	72.92	65.44	68.89
TransECPE(BERT)	88.79	83.15	85.88	**78.74**	66.89	72.33	**77.08**	65.32	70.72
RankCP(BERT)	**91.23**	89.99	**90.57**	74.61	77.88	76.15	71.19	**76.30**	73.60
PTN (BERT)	85.09	91.59	88.19	74.87	**77.90**	**76.31**	76.41	72.40	**74.30**

Table 6 Ablation study of removing CPE, ECP or CSA respectively from the complete model PTN (%). Best results are in bold

Models	Emotion extraction			Cause extraction			Emotion-cause pair extraction		
	P	R	F1	P	R	F1	P	R	F1
PTN	**84.47**	**82.78**	**83.60**	**71.75**	**64.70**	**67.99**	**76.00**	**59.18**	**66.50**
w/o CPE	83.85	81.64	82.71	71.50	64.12	67.57	75.44	58.43	65.83
w/o ECP	83.58	81.87	82.69	70.81	64.09	67.22	75.70	58.11	65.69
w/o CSA	82.55	81.93	82.19	70.96	63.43	66.87	76.94	56.88	65.28

Table 7 Examples of predicted emotion-cause pairs of RankCP and PTN. The clauses in bold and underline respectively denote the emotions and causes

Document	Ground truth	RankCP	PTN
c_1:[*The headmaster persisted in standing on the podium with his stumped legs for* 34 *years,*] c_2:[*and cared about his students when he was critically ill.*] c_3:[***The villagers expressed their highest respect for him.***]	(c_3, c_1) (c_3, c_2)	(c_3, c_2)√	(c_3, c_1)√ (c_3, c_2)√
c_1:[*A group of caring volunteers surrounded the couple of Li Shiming and their two-year-old son Li Muyixin and walked into the house happily.*] c_2:[*For Li Shiming's family*], c_3:[***that day was a happy day***], c_4:[*and the son who had been abducted for* 45 *days finally returned home that day.*]	(c_3, c_4)	(c_3, c_4)√ (c_1, c_3)×	(c_3, c_4)√

of-the-art method RankCP, we present two examples and their predicted results in Table 7.

For the first document, there are two emotion-cause pairs (c_3, c_1) and (c_3, c_2). The former clause is the emotion, and the latter denotes cause. As aforementioned, RankCP regards the predicted top-1 pair as the emotion-cause pair. For the remaining top-(N–1) pairs, if the pair contains sentiment words of the external sentiment lexicon, RankCp extracts it as the emotion-cause pair. It is a pity there is no explicit sentiment word in the first document. Thus RankCP only selects top-1 pair (c_3, c_2) and fails to extract the emotion-cause pair (c_3, c_1). In contrast, our PTN successfully extract the emotion-cause pairs (c_3, c_1) and (c_3, c_2). In the second document example, the clauses c_1 and c_3 contain explicit sentiment words "happily" and "happy". Although the pair (c_1, c_3) is not an emotion-cause pair, RankCP still extracts it as the predicted result according to the external sentiment lexicon.

From the above examples, we can find the performance of RankCP heavily depends on the external sentiment lexicon. In contrast, our method achieves better performance without these constraints.

5.7 The effect of document length on PTN

We present the performance of our model PTN under different ranges of document length (i.e., number of clauses in a document) to investigate the effect of document length on emotion-cause pair extraction. Figure 4 shows the corresponding results. We can observe that the three metrics, including precision, recall, and F1-score, gradually decline with the increase of document length. The results indicate that long documents have a negative effect on emotion-cause pair extraction. Therefore, how to model long-distance dependency effectively is still a challenge for the ECPE task.

6 Conclusions

Existing studies adopt pipelined methods or end-to-end multi-task learning for the ECPE task, thus suffering from error propagation and potential suboptimal results of emotion-cause pair extraction. In this work, we propose a novel tagging framework PTF to address the ECPE task in an end-to-end way. Different from prior works, PTF transforms emotion extraction, cause extraction, and casual detection between

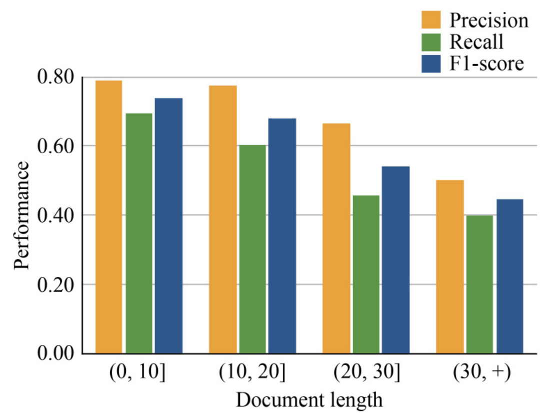

Fig. 4 The performance of PTN on the ECPE task with different document lengths. The proportions of document length ranges (0, 10], (10, 20], (30, 40], (30, +) are respectively 21.53%, 62.56%, 13.33%, and 2.58% in the testing set

emotions and causes into relation classification of clause-pairs, and thus successfully integrates the whole ECPE into one global-optimized and unified tagging task. Based on this novel framework, we design an end-to-end neural network model and introduce three helpful information-enhancing mechanisms to perform the ECPE task. The experimental results and analysis on the benchmark dataset prove the effectiveness of our method.

Acknowledgements This work was supported by the National Natural Science Foundation of China (NSFC) (Grant Nos. 61976114 and 61936012) and the National Key R&D Program of China (2018YFB1005102).

References

1. Fox E. Emotion Science: Cognitive and Neuroscientific Approaches to Understanding Human Emotions. New York: Palgrave Macmillan, 2008
2. Brosch T, Scherer K R, Grandjean D, Sander D. The impact of emotion on perception, attention, memory, and decision-making. Swiss Medical Weekly, 2013, 143: w13786
3. Quan C, Ren F. Construction of a blog emotion corpus for Chinese emotional expression analysis. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing. 2009, 1446–1454
4. Bellegarda J R. Emotion analysis using latent affective folding and embedding. In: Proceedings of the 2010 NAACL HLT Workshop on

Computational Approaches to Analysis and Generation of Emotion in Text. 2010, 1–9
5. Qadir A, Riloff E. Learning emotion indicators from tweets: hashtags, hashtag patterns, and phrases. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. 2014, 1203–1209
6. Lee S Y M, Chen Y, Huang C R. A text-driven rule-based system for emotion cause detection. In: Proceedings of the 2010 NAACL HLT Workshop on Computational Approaches to Analysis and Generation of Emotion in Text. 2010, 45–53
7. Russo I, Caselli T, Rubino F, Boldrini E, Martínez-Barco P. EMOCause: an easy-adaptable approach to extract emotion cause contexts. In: Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis. 2011, 153–160
8. Neviarouskaya A, Aono M. Extracting causes of emotions from text. In: Proceedings of the 6th International Joint Conference on Natural Language Processing. 2013, 932–936
9. Gao K, Xu H, Wang J. A rule-based approach to emotion cause detection for Chinese micro-blogs. Expert Systems with Applications, 2015, 42(9): 4517–4528
10. Gui L, Wu D, Xu R, Lu Q, Zhou Y. Event-driven emotion cause extraction with corpus construction. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016, 1639–1649
11. Cheng X, Chen Y, Cheng B, Li S, Zhou G. An emotion cause corpus for Chinese microblogs with multiple-user structures. ACM Transactions on Asian and Low-Resource Language Information Processing, 2018, 17(1): 6
12. Fan C, Yan H, Du J, Gui L, Bing L, Yang M, Xu R, Mao R. A knowledge regularized hierarchical approach for emotion cause analysis. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. 2019, 5614–5624
13. Xia R, Ding Z. Emotion-cause pair extraction: a new task to emotion analysis in texts. In: Proceedings of the 57th Conference of the Association for Computational Linguistics. 2019, 1003–1012
14. Song H, Zhang C, Li Q, Song D. End-to-end emotion-cause pair extraction via learning to link. 2020, arXiv preprint arXiv: 2002.10710
15. Tang H, Ji D, Zhou Q. Joint multi-level attentional model for emotion detection and emotion-cause pair extraction. Neurocomputing, 2020, 409: 329–340
16. Ding Z, Xia R, Yu J. ECPE-2D: emotion-cause pair extraction based on joint two-dimensional representation, interaction and prediction. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020, 3161–3170
17. Wei P, Zhao J, Mao W. Effective inter-clause modeling for end-to-end emotion-cause pair extraction. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020, 3171–3181
18. Lee S Y M, Chen Y, Huang C R, Li S S. Detecting emotion causes with a linguistic rule-based approach. Computational Intelligence, 2013, 29(3): 390–416
19. Li W, Xu H. Text-based emotion classification using emotion cause extraction. Expert Systems with Applications, 2014, 41(4): 1742–1749
20. Gao K, Xu H, Wang J. Emotion cause detection for Chinese micro-blogs based on ECOCC model. In: Proceedings of the 19th Pacific-Asia Conference on Knowledge Discovery and Data Mining. 2015, 3–14
21. Yada S, Ikeda K, Hoashi K, Kageura K. A bootstrap method for automatic rule acquisition on emotion cause extraction. In: Proceedings of the 2017 IEEE International Conference on Data Mining Workshops. 2017, 414–421
22. Chen Y, Lee S Y M, Li S, Huang C R. Emotion cause detection with linguistic constructions. In: Proceedings of the 23rd International Conference on Computational Linguistics. 2010, 179–187
23. Gui L, Yuan L, Xu R, Liu B, Lu Q, Zhou Y. Emotion cause detection with linguistic construction in Chinese Weibo text. In: Proceedings of the 3rd CCF International Conference on Natural Language Processing and Chinese Computing. 2014, 457–464
24. Ghazi D, Inkpen D, Szpakowicz S. Detecting emotion stimuli in emotion-bearing sentences. In: Proceedings of the 16th International Conference on Computational Linguistics and Intelligent Text Processing. 2015, 152–165
25. Gui L, Xu R, Lu Q, Wu D, Zhou Y. Emotion cause extraction, a challenging task with corpus construction. In: Proceedings of the 5th National Conference on Social Media Processing. 2016, 98–109
26. Xu R, Hu J, Lu Q, Wu D, Gui L. An ensemble approach for emotion cause detection with event extraction and multi-kernel SVMs. Tsinghua Science and Technology, 2017, 22(6): 646–659
27. Gui L, Hu J, He Y, Xu R, Lu Q, Du J. A question answering approach to emotion cause extraction. 2017, arXiv preprint arXiv: 1708.05482
28. Li X, Song K, Feng S, Wang D, Zhang Y. A co-attention neural network model for emotion cause analysis with emotional context awareness. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018, 4752–4757
29. Yu X, Rong W, Zhang Z, Ouyang Y, Xiong Z. Multiple level hierarchical network-based clause selection for emotion cause extraction. IEEE Access, 2019, 7: 9071–9079
30. Ding Z, He H, Zhang M, Xia R. From independent prediction to reordered prediction: integrating relative position and global label information to emotion cause identification. In: Proceedings of the 33rd AAAI Conference on Artificial Intelligence. 2019, 6343–6350
31. Xia R, Zhang M, Ding Z. RTHN: a RNN-transformer hierarchical network for emotion cause extraction. In: Proceedings of the 28th International Joint Conference on Artificial Intelligence. 2019, 5285–5291
32. Wu S, Chen F, Wu F, Huang Y, Li X. A multi-task learning neural network for emotion-cause pair extraction. In: Proceedings of the 24th European Conference on Artificial Intelligence. 2020, 2212–2219
33. Fan C, Yuan C, Gui L, Zhang Y, Xu R. Multi-task sequence tagging for emotion-cause pair extraction via tag distribution refinement. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2021, 29: 2339–2350
34. Cheng Z, Jiang Z, Yin Y, Yu H, Gu Q. A symmetric local search network for emotion-cause pair extraction. In: Proceedings of the 28th International Conference on Computational Linguistics. 2020, 139–149
35. Yu J, Liu W, He Y, Zhang C. A mutually auxiliary multitask model with self-distillation for emotion-cause pair extraction. IEEE Access, 2021, 9: 26811–26821
36. Fan C, Yuan C, Du J, Gui L, Yang M, Xu R. Transition-based directed graph construction for emotion-cause pair extraction. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020, 3707–3717
37. Yuan C, Fan C, Bao J, Xu R. Emotion-cause pair extraction as sequence labeling based on a novel tagging scheme. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing. 2020, 3568–3573
38. Cheng Z, Jiang Z, Yin Y, Li N, Gu Q. A unified target-oriented sequence-to-sequence model for emotion-cause pair extraction. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2021, 29: 2779–2791

39. Yang Z, Yang D, Dyer C, He X, Smola A, Hovy E. Hierarchical attention networks for document classification. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2016, 1480−1489
40. Wu Z, Dai X Y, Yin C, Huang S, Chen J. Improving review representations with user attention and product attention for sentiment classification. In: Proceedings of the 32nd AAAI Conference on Artificial Intelligence. 2018, 5989−5996
41. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation, 1997, 9(8): 1735–1780
42. Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. In: Proceedings of the 3rd International Conference on Learning Representations. 2015
43. Devlin J, Chang M W, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019, 4171−4186
44. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. In: Proceedings of the 26th International Conference on Neural Information Processing Systems. 2013, 3111−3119
45. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 2014, 15(1): 1929–1958
46. Kingma D P, Ba J. Adam: a method for stochastic optimization. In: Proceedings of the 3rd International Conference on Learning Representations. 2015
47. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser Ł, Polosukhin I. Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017, 6000−6010

Zhen Wu is a PhD candidate of the National Key Lab for Novel Software Technology, Department of Computer Science & Technology, Nanjing University, China. He received his BEng degree from Nanjing University of Science and Technology, China in 2016. In the same year, he was admitted to pursue a PhD degree at Nanjing University, China. His research interests include natural language processing and sentiment analysis.

Xinyu Dai is a Professor of the School of Artificial Intelligence, Nanjing University, China. He received his PhD degree in the Department of Computer Science & Technology, Nanjing University, China in 2005. He joined Nanjing University, as an assistant professor in 2005, worked as an associate professor from 2008, and became professor in 2017. His research interests majorly include language processing and intelligence, knowledge engineering, and human-machine communication.

Rui Xia is a Professor of the School of Computer Science and Engineering, Nanjing University of Science and Technology, China. He received the BSc degree from Southeast University, China in 2004, the MSc degree from East China University of Science and Technology, China in 2007, and the PhD degree from the Institute of Automation, Chinese Academy of Sciences, China in 2011. His research interests include natural language processing, machine learning, and data mining.