

Event-Driven Emotion Cause Extraction with Corpus Construction

Lin Gui¹, Dongyin Wu¹, Ruifeng Xu^{1,2*}, Qin Lu³ and Yu Zhou¹

1. School of Computer Science and Technology, Harbin Institute of Technology,
Shenzhen Graduate School, Shenzhen, China

2. Guangdong Provincial Engineering Technology Research Center for Data Science

3. Department of Computing, the Hong Kong Polytechnic University, Hong Kong

guilin.nlp@gmail.com; wudongyinhit@gmail.com; xuruifeng@hitsz.edu.cn;
csluqin@comp.polyu.edu.hk; zhoyu.nlp@gmail.com

Abstract

In this paper, we present our work in emotion cause extraction. Since there is no open dataset available, the lack of annotated resources has limited the research in this area. Thus, we first present a dataset we built using SINA city news. The annotation is based on the scheme of the W3C Emotion Markup Language. Second, we propose a 7-tuple definition to describe emotion cause events. Based on this general definition, we propose a new event-driven emotion cause extraction method using multi-kernel SVMs where a syntactical tree based approach is used to represent events in text. A convolution kernel based multi-kernel SVM are used to extract emotion causes. Because traditional convolution kernels do not use lexical information at the terminal nodes of syntactic trees, we modify the kernel function with a synonym based improvement. Even with very limited training data, we can still extract sufficient features for the task. Evaluations show that our approach achieves 11.6% higher F-measure compared to referenced methods. The contributions of our work include resource construction, concept definition and algorithm development.

1 Introduction

With the rapid growth of Internet, people can easily share experiences and emotions through this powerful medium anywhere and anytime. How to analyze the emotions of individuals through their writings becomes a new challenge for NLP. In recent years, s-

tudies in emotion analysis focus on emotion classification including detection of emotions expressed by writers of text (Gao et al., 2013) as well as prediction of reader emotions (Chang et al., 2015). There are also some information extraction tasks in emotion analysis, such as extracting the feeler of emotion (Das and Bandyopadhyay, 2010). However, these methods need to observe emotion linked expressions. Sometimes, however, we care more about the stimuli, or the cause of an emotion. For instance, manufacturers want to know why people love, or hate a certain product. The White House may also prefer to know the cause of the emotional text “Let us hit the streets” rather than the distribution of different emotions.

There are three main challenges in the study of emotion cause extraction. The first is that, up to now, there is no open dataset available for emotion cause extraction. This may explain why there are only few studies on emotion causes. The second is that, there is no formal definition about event in emotion cause extraction even though some researches claim that they extract events of emotion causes (Lee et al., 2010; Chen et al., 2010). The third is that, due to the complexity in annotation, the size of corpus for emotion cause extraction is usually very small. Due to this limitation, many machine learning methods are not suited for emotion cause detection. How to mine deep knowledge of a language for emotion causes is another thorny issue.

In this paper, we first present an annotated dataset for emotion cause extraction to be released to the public. We then propose to use a 7-tuple to define emotion cause events. Based on this general defi-

*corresponding author

tion, we then present a new event-driven emotion cause extraction method. The basic idea is to extract events in the context of emotional text through dependency parsing. Then, a syntactic structure is used to represent nearby events. Based on this structured representation of events, a modified convolution kernel which also takes lexical features (as terminal nodes) is used to determine whether an event is emotion cause relevant. This method can detect all possible combinations of syntactic structures to obtain sufficient features for emotion analysis using a limited training set. Compared to existing methods, which either use manual rules or commonsense knowledge to extend information, our approach is completely machine learning based and it still achieves state-of-the-art performance. The contributions of this work include both resource development and algorithm development.

The rest of the paper is organized as follows. Section 2 provides a review of related works on emotion analysis. Section 3 presents emotion cause related definitions and the construction of emotion cause extraction corpus. Section 4 gives the event-driven emotion cause extraction method and section 5 is the evaluations and discussions. Section 6 concludes this work and gives the future directions.

2 Related Works

Identifying emotion categories in text is an essential subject in NLP and its applications (Liu, 2015). Moreover, emotion causes can provide important information on why there is any emotion changes. In this section, we introduce related works on the emotion analysis and emotion cause extraction.

The first issue in emotion analysis is to determine the taxonomy of emotions. Researchers have proposed a list of primary emotions (Plutchik, 1980; Ekman, 1984; Turner, 2000). In this study, we adopt Ekman's emotion classification (Ekman, 1984), which identifies six primary emotions, namely *happiness*, *sadness*, *fear*, *anger*, *disgust* and *surprise*, known as the "Big6"¹ scheme in the W3C Emotion Markup Language. This list is agreed upon by most previous works in Chinese emotion analysis.

The second issue is how to do emotion classification and emotion information extraction.

¹<http://www.w3.org/TR/emotion-voc/xml#big6>

Beck (Beck et al., 2014) proposed a Multi-task Gaussian-process based method for emotion classification. Xu (Xu et al., 2012) used a coarse to fine method to classify emotions in Chinese blog. Gao (Gao et al., 2013) proposed a joint model to co-train a polarity classifier and an emotion classifier. Chang (Chang et al., 2015) used linguistic template to predict reader's emotions. Das (Das and Bandyopadhyay, 2010) used an unsupervised method to extract emotion feelers from Bengali blog. There are other studies focused on joint learning with sentiment (Luo et al., 2015; Mohtarami et al., 2013), emotion in tweets or blog (Hasegawa et al., 2013; Qadir and Riloff, 2014; Ou et al., 2014; Liu et al., 2013; Quan and Ren, 2009), and emotional lexicon construction (Yang et al., 2014; Staiano and Guerini, 2014; Mohammad and Turney, 2013). However, these related works all focused on analysis of emotion expressions rather than emotion causes..

Sophia M. Y. Lee first proposed a task on emotion cause extraction (Lee et al., 2010). They manually constructed a corpus from Academia Sinica Balanced Chinese Corpus. Based on this corpus, Chen and Lee (Chen et al., 2010) proposed a rule based method to detect emotion causes. The basic idea is to make linguistic rules for cause extraction. Some studies (Gui et al., 2014; Li and Xu, 2014; Gao et al., 2015) extended the rule based method to informal text in Weibo text (Chinese tweets).

Other than rule based methods, Ghazi (Ghazi et al., 2015) used CRFs to extract emotion causes. However, it requires emotion cause and emotion keywords to be in the same sentence. I. Russo (Russo et al., 2011) proposed a crowd-sourcing method to obtain emotion cause related commonsense knowledge. But it is challenging to extend the commonsense knowledgebase automatically.

Resources used in the above works are not publicly accessible. Most of the methods used are rule based. Learning based methods are quite limited because annotated data is quite small in size due to high cost for annotation. Thus, rule based methods seem to be the easiest way to achieve acceptable performance. Since machine learning methods require more knowledge, which is difficult to generalize. So automatic methods only focused on simple text genre.

3 Construction of Emotion Cause Corpus

In this section, we first describe the linguistic phenomenon in emotion expressions. It serves as the inspiration to develop the annotated dataset. We then introduce details of the annotation scheme, followed by the construction of the dataset.

3.1 Linguistic Phenomenon of Emotion Causes

Emotion causes play an important role in emotion expressions. An emotion cause reveals the stimulus of an emotion. Considering linguistic phenomenon of emotion causes, we follow three basic principles in corpus construction: (1) Keep the whole context of emotion expression; (2) The basic processing unit is at the clause level; and (3) Use of formal text.

In written text, there is an emotion keyword, which is used to express an emotion, in the context of the emotion cause. Thus, finding the appropriate context of emotion keywords in the annotation is the pre-requisite to identify its cause. It is the reason why we keep the whole context of emotion keywords.

Another important kind of cues is the presence of conjunctions and prepositions. These words indicate the discourse information between clauses. In order to make use of discourse information, the basic analysis unit should be at clause level rather than at sentence level.

In the third principle, we choose the formal text in corpus construction. According to the related works, emotion expressions can have overlapping emotion cause and emotion target (Gui et al., 2014) in informal text. This is why some studies even incorporate cause extraction with target identification to improve performance. However, our focus is on emotion cause identification. We use formal news text to avoid the potential mix up.

3.2 Collection and Annotation

We first take 3 years (2013-15) Chinese city news from NEWS SINA² containing 20,000 articles as the raw corpus. Based on a list of 10,259 Chinese primary emotion keywords (**keywords** for short) (Xu et al., 2008), we extract 15,687 instances by keyword matching from the raw data. Here, we call the presence of an emotion keyword as an **instance** in

the corpus. For each matched keyword, we extract three preceding clauses and three following clauses as the context of an instance. If a sentence has more than 3 clauses in each direction, the context will include the rest of the sentence to make the context complete. For simplicity, we omit cross paragraph context.

Note that the presence of keywords does not necessarily convey emotional information due to different possible reasons such as negative polarity and sense ambiguity. For example, “祝愿/wishes” is an emotion word of “happiness”. It can also be the name of a song. Also, the presence of emotion keywords does not necessarily guarantee the existence of emotional cause neither. After removing those irrelevant instances, there are 2,105 instances remain. For each emotional instance, two annotators manually annotate the emotion categories and the cause(es) in the W3C Emotion Markup Language (EML) format. Ex1 shows an example of an annotated emotional sentence in the corpus, presented by the original simplified Chinese, followed by its English translation. To save space, we remove the xml tags in the annotation. The original annotated data is in a subsidiary file³. The basic analysis unit is a clause. Emotion cause is marked by <cause>, and the emotion keyword is marked by <keywords>. Emotion type, POS, position and the length of annotation are also annotated in Emotionml format.

Ex.1: 朱某今年55岁，1979年参加工作时才19岁，已有36年的手艺。“我当时被分到丹阳南京理发店工作，这是当时丹阳最大的理发店。我在那儿获得了好多证书和荣誉。” <cause POS=“v” Dis=“-1”>说起自己的荣誉</cause>，朱某很是<keywords type=happiness>自豪</keywords>。

Mr. Zhu is 55 years old. He started working in 1979 as a barber when he was 19, and has 36 years of experience. “I was assigned to work at the Barbershop in Danyang, Nanjing. It is the largest barbershop in Danyang. I won many awards and honors there.” <cause POS=“v” Dis=“-1”>Talking about his honors</cause>, Mr. Zhu is so <keywords type=“happiness”> proud </keywords>.

Ex.1 only contains one cause. However, one keyword may have more than one corresponding emotion causes. In Ex.2, there are two relevant causes

²<http://news.sina.com.cn/society/>

³http://hlt.hitsz.edu.cn/?page_id=694

Item	Number
Instance	2,105
Clauses	11,799
Emotion Cause	2,167
Document with 1 emotion	2,046
Document with 2 emotion	56
Document with 3 emotion	3

Table 1: Details of the Dataset

for one keyword. In our dataset, only 59 instances have two or more causes.

Ex.2: 劝说过程中, 消防官兵了解到, 该女子是由于<cause POS="v" Dis="-2">对方拖欠工程款</cause>, <cause POS="v" Dis="-1">家中又急需用钱</cause>, <keywords type=sadness>无奈</keywords>才选择跳楼轻生。

During persuasion, firemen realized that the woman attempted suicide because of <cause POS="v" Dis="-2">the hold back of wages by the employer</cause>, and <cause POS="v" Dis="-1">her family asked for money urgently</cause>, she feels <keywords type=sadness>helpless</keywords> and thus

3.3 Details of Dataset and Its Annotations

Each instance in our dataset contains only one emotion keyword and at least one emotion cause. It is ensured that the keyword instance and the causes are relevant. The number of extracted instances, clauses, and emotion causes are listed in Table 1. Note that 97.2% of the instances has only one emotion cause, and instances that have two and three emotion causes hold 2.6% and 0.2% respectively. Table 2 shows the distribution of emotion types and Table 3 shows the distribution of cause positions. In the latter we can see that 78% emotion causes adjoin the emotion keywords at the clause level. Apparently, position plays a very important role in emotion cause extraction. Thus, using distance based features for emotion cause extraction is rational and necessary. Table 4 lists the phrase types of emotion causes. Verbs and verb phrases cover 93% of all cause events. Thus, our learning algorithm mainly focus on them.

Two annotators work independently during the annotation process. The key point is to distinguish clause level and phrase level in cause annotation. The clause level labels the clause which contains the emotion cause. The phrase level determines the boundary of an emotion cause. When two annota-

Emotion	Number	Percentage
Happiness	544	25.83%
Sadness	567	26.94%
Fear	379	18.00%
Anger	302	14.35%
Disgust	225	10.69%
Surprise	88	4.18%

Table 2: Distribution of Emotion Types

Position	Number	Percentage
Previous 3 clauses	37	1.71%
Previous 2 clauses	167	7.71%
Previous 1 clauses	1,180	54.45%
In the same clauses	511	23.58%
Next 1 clauses	162	7.47%
Next 2 clauses	48	2.22%
Next 3 clauses	11	0.51%
Other	42	1.94%

Table 3: Cause Position of Each Emotion

tors have different opinion on one instance at clause level, we involve a third annotator as the arbitrator. In the phrase level, we use the larger boundary of the two annotations when they have the same annotation at the clause level. We reach 0.9287 for the kappa value on clause level annotation which confirmed the reliability of our annotation.

4 Event-Driven Emotion Cause Extraction

Due to the complexity of annotation in emotion cause identification, the size of annotated corpus is usually small. Since we aim to use machine learning methods to automatically learn and identify causes, we use a convolution kernel to detect all possible combinations in the syntactic structure. This allows learning from syntactic representations for emotion cause extraction. The basic idea of our proposed method is to use a tree-structure representation to capture features for emotion cause identification. For training data, we extract all valid tree structures for each event, referred to as the ETs (Event Trees). If an event is a cause, the corresponding ET is positive. Otherwise, the corresponding ET is negative. Then, we train a convolution kernel and a

POS/phrase type	Number	Percentage
Noun/Noun phrase	147	6.78%
Verb/Verb phrase	2020	93.21%

Table 4: Distribution of the POS Tag

multi-kernel SVMs using the training set to classify candidate ETs in the testing set. Since more than 97% emotion keywords only have one cause, and more than 95% causes are near the emotion keywords, candidate ETs are extracted from the context of emotion keywords. We only choose the ET with the highest probability in the classification result as the emotion cause.

4.1 Event Tree Construction

Even though there are related works on event identification in emotion cause detection, there is no formal definition of events in area of artificial intelligence (AI), researchers, such as Radinsky (Radinsky et al., 2012), gave a formal definition of an event as “action, actor, object, instrument, location and time”. In our work, we need to give clear definition of event first.

In emotion cause extraction, the components of an event should be simpler. We are only interested in the action, the actor and the object, which are denoted as P , O_1 , O_2 , respectively, following the conventions in AI. Since Chinese is a SVO language, the actor is the subject and the action is the verb. The subject and the object of a sentence may have attributes and a predicate may have adverbial and complement. Since these components may also be helpful in emotion cause extraction, we formally define an emotion cause event as a 7-tuple:

$$e = (Att_{O_1}, O_1, Adv, P, Cpl, Att_{O_2}, O_2).$$

Here, Att_{O_1} is the attribute of O_1 ; Att_{O_2} is the attribute of O_2 ; Adv is the adverbial of the predicate P ; and Cpl is P 's complement. In case syntactic components are not present, NIL values are used. Note that the main cue in an event is P , the action. So, in our algorithm, we extract all verbs from the text, and use dependency parsing⁴ to extract all relevant syntactic components specified in e . Then, we can construct an ET.

An ET has a fixed height of four levels. The top level is the root node. Since Chinese is a SVO language, the descendant of the root is S(subject), V(verb), and O(object). Then, the seven event components can be categorized and filled up in the relevant slots. (Att_{O_1}, O_1) belongs to $S(O_1)$, (Adv, P, Cpl) belong to V , and (Att_{O_2}, O_2) belongs

⁴<https://github.com/HIT-SCIR/ltp>

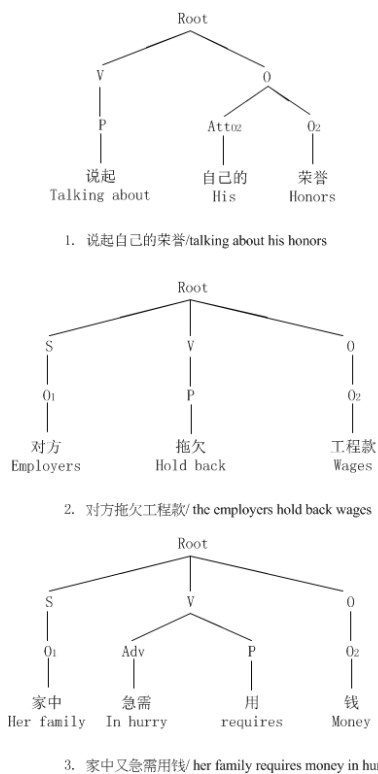


Figure 1: Example ETs of Emotion Causes.

to O . Then we can get the ET based on the definition of an event.

Let us review Ex.1 and Ex.2 again. There are three emotion cause events below with their corresponding ETs shown in Figure 1.

1. “说起自己的荣誉/Talking about his honors”
2. “对方拖欠工程款/ the hold back wages by employers”
3. “家中又急需用钱/ her family asked for money urgently”

After the construction of the ETs, emotion cause extraction becomes a classification problem. If an ET is an emotion cause, the label should be positive. Otherwise, the label should be negative. A binary classifier should be used.

4.2 Emotion Cause Extraction

After the construction of ETs, we obtain positive and negative ET samples. Due to small amount of training samples, it is necessary to capture all features in the ETs. We choose convolution kernel based SVMs because it can search all possible syntactic features under a tree structure.

Convolution kernel function

The convolution kernel, also known as the tree kernel (Collins and Duffy, 2002), is widely used in many NLP tasks (Srivastava et al., 2013; Moschitti, 2006). For any two inputs T_1 and T_2 based on a tree structure, the kernel is defined as:

$$K(T_1, T_2) = \sum_{n_1 \in T_1} \sum_{n_2 \in T_2} \delta(n_1, n_2). \quad (1)$$

Here, n_1 and n_2 are tree nodes. δ is a function defined recursively:

1. $\delta(n_1, n_2) = 0$ if the productions of n_1 and n_2 are different; 2. Else, $\delta(n_1, n_2) = 1$ if n_1 and n_2 are matching in pre-terminals; 3. Otherwise,

$$\delta(n_1, n_2) = \prod_i (1 + \delta(c(n_1, i), c(n_2, i))).$$

Here, $c(n, i)$ is the i -th node of n .

However, the above tree kernel definition does not consider terminals, which means that the actual words in a sentence are ignored. As emotions causes are semantically meaningful, we need to incorporate lexical information into the convolution kernel.

Modified kernel function

In order to distinguish different ETs, we need to modify the definition of the tree kernel to include lexical words in a clause. So we add one more definition to include the terminals:

4. If n_1 and n_2 are terminal nodes, $\delta(n_1, n_2) = 1$ if and only if n_1 and n_2 are synonyms. Otherwise $\delta(n_1, n_2) = 0$.

Here a synonym is defined in Tongyici Cilin (Extended).⁵ which has 17,817 synonyms and 77,343 words. We use the synonym rather than word matching because the size of the corpus is limited. simple word matching is quite sparse.

Let K_{ET-O} denote the original kernel and K_{ET-M} denote the modified kernel, respectively. It can be easily proven that K_{ET-M} is a valid kernel function. Following the notation in (Collins and Duffy, 2002), $K_{ET-O} = \sum_i h_i(T_1) \cdot h_i(T_2)$, where $h_i(T_1) = \sum_{n_1 \in N_1} I_i(n_1)$, $h_i(T_2) = \sum_{n_2 \in N_2} I_i(n_2)$ and the function $I_i(n)$ is 1 if the sub-tree i is rooted at node n and 0 otherwise. So the original tree kernel is an inner product and the kernel matrix is

semi-definite. In our modified kernel, the function $I_i(n)$ is more complicated. Beside the definition above, it has the following additional definition: $I_i(n)$ is 1 if i is a terminal node and it is a synonym of n . The new indicator is marked as $I'_i(n)$. Then we have: $K_{ET-M}(T_1, T_2) = \sum_{n_1 \in T_1} \sum_{n_2 \in T_2} \sum_i I'_i(n_1) I'_i(n_2)$. This means that the modified kernel is symmetrical and the kernel matrix is semi-definite. In our work, K_{ET-M} uses SVM optimization and the code is from SVM-light-TK⁶.

Multi-kernel function

Since there are only syntactic information and synonyms in the convolution kernel based method, we need to add some lexical features. Given a 7-tuple event e , we obtain the bag-of-words based or word embedding based representation for each component in e , and the distance between a component and emotion keywords are used as the features, respectively. Let the features of each component in e be R_i , for every $i \in e$. Then, we can capture the feature set, F , of an ET by a joint operation, called the ET features:

$$F = \{R_{AttO_1} \oplus R_{O_1} \oplus \dots \oplus R_{AttO_2}\}. \quad (2)$$

We can join the ET features with syntactic information by a multi-kernel function. For any two ETs T_1 and T_2 , with the respective features F_1 and F_2 , the two new multi-kernels can be defined as:

$$K_{new+O}(T_1, T_2) = K_{ET-O}(T_1, T_2) + K_{vec}(F_1, F_2), \quad (3)$$

$$K_{new*O}(T_1, T_2) = K_{ET-O}(T_1, T_2) \times K_{vec}(F_1, F_2), \quad (4)$$

$$K_{new+M}(T_1, T_2) = K_{ET-M}(T_1, T_2) + K_{vec}(F_1, F_2), \quad (5)$$

$$K_{new*M}(T_1, T_2) = K_{ET-M}(T_1, T_2) \times K_{vec}(F_1, F_2). \quad (6)$$

Here, K_{vec} denotes a kernel function which can be a linear kernel, a polynomial kernel or a Gaussian kernel. The next step is to train the classifier based on the multi-kernel function.

The training data is already in labeled ET format. To prepare testing data, we extract all ETs from a given instance as candidate ETs. A classifier is used to obtain the probability of emotion cause for each ET to produce a ranked list of candidate ETs. The ET with the highest rank serves as the cause event for the current instance.

⁵http://ir.hit.edu.cn/demo/ltp/Sharing_Plan.htm

⁶<http://disi.unitn.it/moschitti/Tree-Kernel.htm>

5 Performance Evaluations

5.1 Experimental Setup

In the experiments, we stochastically select 90% of the dataset as training data and 10% as testing data. In order to obtain statistically credible results, we evaluate our methods and the reference methods 25 times. We conduct two sets of experiments. The first one evaluates the performance at the clause level to identify the clauses that contain emotion causes. The second one evaluates emotion causes using verb classification. This is because 93.21% of emotion causes are verb/verb phrase and verbs serve as the action component in event definition.

5.2 Emotion Cause Extraction

We use the commonly accepted measure proposed by Lee (Lee et al., 2010) for emotion cause extraction (Gao et al., 2015; Li and Xu, 2014). In this measure, if a proposed emotion cause covers the annotated answer, the sequence is considered correct. The precision, recall, and F-measure are defined by

$$Precision = \frac{\sum correct_cause1}{\sum proposed_cause1},$$

$$Recall = \frac{\sum correct_cause1}{\sum annotated_cause1},$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall}.$$

In the experiment, evaluation is conducted for the following works:

- 1.**RB**(Rule based method): Among several rule based methods (Lee et al., 2010; Gui et al., 2014; Li and Xu, 2014). We use lee2010’s rules (listed in **Appendix** of this paper).
- 2.**CB**(Commonsense based method): In order to reproduce this method (Russo et al., 2011), we use the Chinese Emotion Cognition Lexicon (Xu et al., 2013) as the commonsense. The lexicon contains more than 5,000 emotion stimulations and their corresponding reflection words.
- 3.**ML**(Rule base features for machine learning): Rules are used as features with other manual features for emotion cause classification (Chen et al., 2010).
4. K_{vec} : Features are defined in Formula (2) in the training of classifier.

Method	Precision	Recall	F-measure
RB	0.6747	0.4287	0.5243
CB	0.2672	0.7130	0.3887
RB+CB	0.5435	0.5307	0.5370
RB+CB+ML	0.5921	0.5307	0.5597
K_{vec}	0.4200	0.4375	0.4285
$K_{word2vec}$	0.4301	0.4233	0.4136
K_{ET-O}	0.3982	0.4134	0.4057
K_{ET-M}	0.4583	0.4745	0.4662
K_{new+O}	0.6446	0.6779	0.6608
K_{new*O}	0.6492	0.6701	0.6595
K_{new+M}	0.6588	0.6927	0.6752
K_{new*M}	0.6673	0.6841	0.6756

Table 5: Performance on the Dataset

5. $K_{word2vec}$: Word2vec (Mikolov et al., 2013) is used to learn word representation. Use the representation according to Formula (2) in the training of classifier.

6. K_{ET-O} : This is the original tree kernel.

7. K_{ET-M} : This is the modified tree kernel in Formula (1).

8. K_{new+O} , K_{new*O} , K_{new+M} and K_{new*M} : Use the multi-kernel gives by formulas from (3) to (6).

The performance result is given in Table 5. Among all methods, K_{new*M} achieves the top performance in F-measure. Compared to other methods, the improvement is significant with p-value less than 0.01 in *t*-test.

Even though RB achieves the top precision, its F-measure is limited by the low recall. Since CB is opposite to RB, the performance by RB+CB is improved. However, the improvement is quite limited, at 0.0127 in F-measure. The F-measure of our reproduced RB is similar to mentioned result of other references (Gui et al., 2014; Li and Xu, 2014). They repeat Lee’s (Lee et al., 2010) method and achieve the F-measure with 0.55 more or less.

(Chen et al., 2010) reported that by using hand-crafted rules as features to train a classifier with some additional features such as conjunction, action and epistemic verbs, performance can be improved significantly. In our experiment, the result is opposite to this claim. The main reason is the samples in (Chen et al., 2010) are less complex. About 85% of the emotion causes are in the same clause where the emotion keywords are. Our corpus is quite different. The percentage of causes in the same clause where the emotion keyword itself is has only about

23.6%. (Chen et al., 2010)’s method does not handle long distance relations well. This explains why it does not work well for our dataset. Although (RB+CB+ML) does not perform well, there is still 0.0334 improvement in F-measure compare to RB. Among our proposed methods, K_{vec} on the ET feature achieves 0.4285 in F-measure. Compare to CB and ML, the performance is not satisfactory. However, as a simple feature to represent lexical information, the performance is acceptable. word2vec also yield similar result. Maybe the joint operation is too simple to handle composition.

For the modified tree kernel K_{ET-M} , the performance is 0.0605 higher than the original tree kernel K_{ET-O} in F-measure. It means that the consideration of terminal node improves the performance of the tree kernel significantly. The modified tree kernel K_{ET-M} is also 0.0377 higher than K_{vec} , and 0.0526 higher than $K_{word2vec}$ in F-measure. This means kernel based syntactic representation does have better generalization ability. The original kernel function K_{ET-O} has syntactic information but no lexicon, and it not only underperforms compared to K_{ET-M} but also K_{vec} and $K_{word2vec}$. This demonstrates our modified kernel function can effectively turn an inferior method into a superior one. Compared to rule based method, the performance still needs to be enhanced and a multi-kernel is necessary. After the combination with ET feature using a multi-kernel, the performance of K_{new*M} achieves a higher level with 0.6756 in F-measure. Compare to RB, the improvement in F-measure is 0.1513. Compare to the combination of existing methods, the improvement is 0.1159. The reason is that our method represents events at the syntactic level. Synonym information gives the model more generalization ability.

5.3 Verb Classification for Emotion Cause

In this section, we examine the performance of ETs classification with respect to verbs identified in the emotion clauses.

ETs Classification

Our method is based on ETs classification to choose the candidate ET with the highest probability. The performance is measured by the verbs in the identified ET. Results are shown in Table 6.

Note that $K_{word2vec}$ performs much better than

Method	Precision	Recall	F-measure
K_{vec}	0.3500	0.2951	0.3192
$K_{word2vec}$	0.3200	0.4833	0.3848
K_{ET-O}	0.3906	0.2773	0.3228
K_{ET-M}	0.3978	0.3303	0.3473
K_{new+O}	0.4211	0.7219	0.5319
K_{new*O}	0.4197	0.7305	0.5331
K_{new+M}	0.4407	0.7694	0.5651
K_{new*M}	0.4532	0.7504	0.5646

Table 6: Performance on ETs Classification

K_{vec} in verb identification, contrary to their similar performance in clause identification. The reason is that extraction result is based on ranking and only top ranked event affects the performance. In other words, precision is more important than recall here. For the same reason, K_{new+M} is better than K_{new*M} in classification of ETs, although only marginally. Nonetheless, using revised convolution kernel with multi-kernel training is still significantly better than the original kernel K_{new*M} which achieves the best performance in Table 5. When the precision of the two methods are similar, such as K_{ET-O} and K_{ET-M} , the effect of recall becomes important.

The multi-kernel not only achieves the best performance on both precision and recall, the increase in performance is also significant with at least 0.2173 (between K_{ET-M} and K_{new*M}). Obviously, multi-kernel is not just a simple voting or joint for the components, it benefits from two kernels to achieve better performance.

5.4 Error Analysis

There are mainly three types of errors in our model. We use case examples to show them.

a) Cascading Events

In some cases, events may happen like a chain reaction. An event that leads to an emotion may be the consequent of another event. Identifying the right event in a chain is more challenging. In the following example:

Ex.3: 约兰·沃森坠入冰冷的水中。<cause>刺骨的冰水</cause>让他感到极其寒冷与<keywords>害怕</keywords>, 约兰·沃森慌忙用不太流利的中文大呼“救命”。

John Watson fell into icy water. <cause>The chilly water</cause> made him feel so cold and <keywords>scared</keywords> John Watson had to

use his broken Chinese to call for help.

the emotion cause should be “刺骨的冰水/*the chilly water*”. Our method output “坠入冰冷的水中/*fell into icy water*” as the emotion cause with probability 60.83%. The probability of the correct cause is 58.89%. As a probability based method, our method does not have the ability to analyze the sequence of events nor the relation between them.

b) Sensory verbs

Sensory verbs usually indicate the emotion cause. There are exceptional cases as shown below:

Ex.4: 了解霸凌事件后。教务主任说，这三名学生知道错了也感到很<keywords>害怕</keywords>，他们<cause>可能面临劳动服务</cause>

After investigation on bullying, the head says that the students realized their mistake and were also <keywords>scared</keywords>. They <cause> may need to do community service</cause>

In this case, the cause of “scared” is the punishment of community service. But the template of “知道…感到/*realized ... and felt*” usually indicate that there is an emotion cause between the two sensory verbs. Our algorithm gives “知道错了/*realized their mistake*” a probability of 61.65% as a cause, although this is incorrect. But, this actually indicates that our method can learn latent patterns in text.

c) Coverage of cause candidates

In the construction of ETs, we use actions as the cue to construct candidate events. However, 6.78% of our clauses do not have action words. So, these clauses are not selected as candidates.

6 Conclusion

In this paper, we present our work on emotion cause extraction. Due to the lack of open resources for this area of study, we first construct an annotated dataset from news text which will be released for public use. We also propose an event-driven emotion cause extraction method to capture the triggering events emotion changes. In this method, we propose a 7-tuple representation of events using syntactic structures to identify events. Based on this structured representation of events and the inclusion of lexical features, a convolution kernel based learning method is designed to train a multi-kernel classifier to identify emotion cause events. Compared to manually constructed rules and commonsense knowledge based

methods, our proposed model can automatically obtain structure features and lexical features to achieve state-of-the-art performance on this dataset.

Acknowledgment

This work was supported by the National Natural Science Foundation of China 61370165, 61632011, National 863 Program of China 2015AA015405, Shenzhen Peacock Plan Research Grant KQCX20140521144507925 and Shenzhen Foundational Research Funding J-CYJ20150625142543470, Guangdong Provincial Engineering Technology Research Center for Data Science 2016KF09. The project is also partially supported by HK GRF grant PolyU 15211/14E.

Appendix

No.	Rules
1	i) C(B/F) + I(F) + E(F) + K(F) ii) E = the nearest Na/Nb/Nc/Nh after I in F iii) C = the nearest (N)+(V)+(N) before I in F/B
2	i) E(B/F) + II/IV/V/VI(B/F) + C(B/F) + K(F) ii) E = the nearest Na/Nb/Nc/Nh before II/IV/V/VI in B/F iii) C = the nearest (N)+(V)+(N) before K in F
3	i) II/IV/V/VI (B) + C(B) + E(F) + K(F) ii) E = the nearest Na/Nb/Nc/Nh before K in F iii) C = the nearest (N)+(V)+(N) after II/IV/V/VI in B
4	i) E(B/F) + K(F) + IV/VII(F) + C(F/A) ii) E = a: the nearest Na/Nb/Nc/Nh before K in F; b: the first Na/Nb/Nc/Nh in B iii) C = the nearest (N)+(V)+(N) after IV/VII in F/A
5	i) E(F)+K(F)+VI(A)+C(A) ii) E = the nearest Na/Nb/Nc/Nh before K in F iii) C = the nearest (N)+(V)+(N) after VI in A
6	i) I(F) + E(F) + K(F) + C(F/A) ii) E = the nearest Na/Nb/Nc/Nh after I in F iii) C = the nearest (N)+(V)+(N) after K in F or A
7	i) E(B/F) + yue4 C yue4 K “the more C the more K” (F) ii) E = the nearest Na/Nb/Nc/Nh before the first yue4 in B/F iii) C = the V in between the two yue4’ s in F
8	i) E(F) + K(F) + C(F) ii) E = the nearest Na/Nb/Nc/Nh before K in F iii) C = the nearest (N)+(V)+(N) after K in F
9	i) E(F) + IV(F) + K(F) ii) E = the nearest Na/Nb/Nc/Nh before IV in F iii) C = IV+(an aspectual marker) in F
10	i) K(F) + E(F) + de “possession”(F) + C(F) ii) E = the nearest Na/Nb/Nc/Nh after K in F iii) C = the nearest (N)+V+(N)+“的”+N after de in F
11	i) C(F) + K(F) + E(F) ii) E = the nearest Na/Nb/Nc/Nh after K in F iii) C = the nearest (N)+(V)+(N) before K in F
12	i) E(B) + K(B) + III (B) + C(F) ii) E = the nearest Na/Nb/Nc/Nh before K in F iii) C = the nearest (N)+(V)+(N) after III in F
13	i) III(B) + C(B) + E(F) + K(F) ii) E = the nearest Na/Nb/Nc/Nh before K in F iii) C = the nearest (N)+(V)+(N) after III in B
14	i) C(B) + E(F) + K(F) ii) E = the nearest Na/Nb/Nc/Nh before K in F iii) C = the nearest (N)+(V)+(N) before K in B
15	i) E(B) + C(B) + K(F) ii) E = the first Na/Nb/Nc/Nh in B iii) C = the nearest (N)+(V)+(N) before K in B

Table 7: Linguistic Rules in RB

Here, C = Cause event; E = Experiencer; K = Keyword/emotion verb; B = Clause before the focus clause; F = Focus clause/the clause containing the emotion verb; A = Clause after the focus clause; I to VII are cue words in (Lee et al., 2010); Na/Nb/Nc/Nh is common noun/proper noun/place noun/pronoun.

References

- Daniel Beck, Trevor Cohn, and Lucia Specia. 2014. Joint emotion analysis via multi-task gaussian processes. In *EMNLP*, pages 1798–1803.
- Yung-Chun Chang, Cen-Chieh Chen, Yu-Lun Hsieh, and WL Hsu. 2015. Linguistic template extraction for recognizing reader-emotion and emotional resonance writing assistance. *ACL-IJCNLP*, pages 775–780.
- Ying Chen, Sophia Yat Mei Lee, Shoushan Li, and Chu-Ren Huang. 2010. Emotion cause detection with linguistic constructions. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 179–187. Association for Computational Linguistics.
- Michael Collins and Nigel Duffy. 2002. New ranking algorithms for parsing and tagging: Kernels over discrete structures, and the voted perceptron. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 263–270. Association for Computational Linguistics.
- Dipankar Das and Sivaji Bandyopadhyay. 2010. Finding emotion holder from bengali blog texts—an unsupervised syntactic approach. In *PACLIC*, pages 621–628.
- Paul Ekman. 1984. Expression and the nature of emotion. *Approaches to emotion*, 3:19–344.
- Wei Gao, Shoushan Li, Sophia Yat Mei Lee, Guodong Zhou, and Chu-Ren Huang. 2013. Joint learning on sentiment and emotion classification. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*, pages 1505–1508. ACM.
- Kai Gao, Hua Xu, and Jiushuo Wang. 2015. A rule-based approach to emotion cause detection for chinese micro-blogs. *Expert Systems with Applications*, 42(9):4517–4528.
- Diman Ghazi, Diana Inkpen, and Stan Szpakowicz. 2015. Detecting emotion stimuli in emotion-bearing sentences. In *Computational Linguistics and Intelligent Text Processing*, pages 152–165. Springer.
- Lin Gui, Li Yuan, Ruifeng Xu, Bin Liu, Qin Lu, and Yu Zhou. 2014. Emotion cause detection with linguistic construction in chinese weibo text. In *Natural Language Processing and Chinese Computing*, pages 457–464. Springer.
- Takayuki Hasegawa, Nobuhiro Kaji, Naoki Yoshinaga, and Masashi Toyoda. 2013. Predicting and eliciting addressee’s emotion in online dialogue. In *ACL (1)*, pages 964–972.
- Sophia Yat Mei Lee, Ying Chen, and Chu-Ren Huang. 2010. A text-driven rule-based system for emotion cause detection. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 45–53. Association for Computational Linguistics.
- Weiyuan Li and Hua Xu. 2014. Text-based emotion classification using emotion cause extraction. *Expert Systems with Applications*, 41(4):1742–1749.
- Huanhuan Liu, Shoushan Li, Guodong Zhou, Chu-Ren Huang, and Peifeng Li. 2013. Joint modeling of news reader’s and comment writer’s emotions. In *ACL (2)*, pages 511–515.
- Bing Liu. 2015. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.
- Kun-Hu Luo, Zhi-Hong Deng, Liang-Chen Wei, and Hongliang Yu. 2015. Jeam: A novel model for cross-domain sentiment classification based on emotion analysis. In *EMNLP*, pages 2503–2508.
- Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119.
- Saif M Mohammad and Peter D Turney. 2013. Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3):436–465.
- Mitra Mohtarami, Man Lan, and Chew Lim Tan. 2013. Probabilistic sense sentiment similarity through hidden emotions. In *ACL (1)*, pages 983–992.
- Alessandro Moschitti. 2006. Efficient convolution kernels for dependency and constituent syntactic trees. In *Machine Learning: ECML 2006*, pages 318–329. Springer.
- Gaoyan Ou, Wei Chen, Tengjiao Wang, Zhongyu Wei, Binyang Li, Dongqing Yang, and Kam-Fai Wong. 2014. Exploiting community emotion for microblog event detection. In *EMNLP*, pages 1159–1168.
- Robert Plutchik. 1980. Emotion: A psychoevolutionary synthesis.
- Ashequl Qadir and Ellen Riloff. 2014. Learning emotion indicators from tweets: Hashtags, hashtag patterns, and phrases. In *EMNLP*, pages 1203–1209.
- Changqin Quan and Fuji Ren. 2009. Construction of a blog emotion corpus for chinese emotional expression analysis. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 3-Volume 3*, pages 1446–1454. Association for Computational Linguistics.
- Kira Radinsky, Sagie Davidovich, and Shaul Markovitch. 2012. Learning to predict from textual data. *Journal of Artificial Intelligence Research*, pages 641–684.
- Irene Russo, Tommaso Caselli, Francesco Rubino, Ester Boldrini, and Patricio Martínez-Barco. 2011. Emocause: an easy-adaptable approach to emotion cause contexts. In *Proceedings of the 2nd Workshop on*

- Computational Approaches to Subjectivity and Sentiment Analysis*, pages 153–160. Association for Computational Linguistics.
- Shashank Srivastava, Dirk Hovy, and Eduard H Hovy. 2013. A walk-based semantically enriched tree kernel over distributed word representations. In *EMNLP*, pages 1411–1416.
- Jacopo Staiano and Marco Guerini. 2014. Depechemood: a lexicon for emotion analysis from crowd-annotated news. *arXiv preprint arXiv:1405.1605*.
- Jonathan H Turner. 2000. *On the origins of human emotions: A sociological inquiry into the evolution of human affect*. Stanford University Press Stanford, CA.
- Linhong Xu, Hongfei Lin, Yu Pan, Hui Ren, and Jianmei Chen. 2008. Constructing the affective lexicon ontology. *Journal of the China Society for Scientific and Technical Information*, 27(2):180–185.
- Jun Xu, Ruifeng Xu, Qin Lu, and Xiaolong Wang. 2012. Coarse-to-fine sentence-level emotion classification based on the intra-sentence features and sentential context. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 2455–2458. ACM.
- Ruifeng Xu, Chengtian Zou, Yanzhen Zheng, Xu Jun, Lin Gui, Bin Liu, and Xiaolong Wang. 2013. A new emotion dictionary based on the distinguish of emotion expression and emotion cognition. *Journal of Chinese Information Processing*, 27(6):82–90.
- Min Yang, Dingju Zhu, and Kam-Pui Chow. 2014. A topic model for building fine-grained domain-specific emotion lexicon. In *ACL (2)*, pages 421–426.