An Approach of Fuzzy Relation Equation and Fuzzy-Rough Set for Multi-label Emotion Intensity Analysis

Chu Wang[1], Daling Wang[1,2(✉)], Shi Feng[1,2], and Yifei Zhang[1,2]

[1] School of Computer Science and Engineering, Northeastern University, Shenyang, China
wangchu@research.neu.edu.cn,
{wangdaling, fengshi, zhangyifei}@cse.neu.edu.cn

[2] Key Laboratory of Medical Image Computing (Northeastern University), Ministry of Education, Shenyang 110819, People's Republic of China

Abstract. There are a large number of subjective texts which contain people's all kinds of sentiments and emotions in social media. Analyzing the sentiments and predicting the emotional expressions of human beings have been widely studied in academic communities and applied in commercial systems. However, most of the existing methods focus on single-label sentiment analysis, which means that only an exclusive sentiment orientation (negative, positive or neutral) or an emotion state (joy, hate, love, sorrow, anxiety, surprise, anger, or expect) is considered for a document. In fact, multiple emotions may be widely coexisting in one document, paragraph, or even sentence. Moreover, different words can express different emotion intensities in the text. In this paper, we propose an approach that combining fuzzy relation equation with fuzzy-rough set for solving the multi-label emotion intensity analysis problem. We first get the fuzzy emotion intensity of every sentiment word by solving a fuzzy relation equation, and then utilize an improved fuzzy-rough set method to predict emotion intensity for sentences, paragraphs, and documents. Compared with previous work, our proposed algorithm can simultaneously model the multi-labeled emotions and their corresponding intensities in social media. Experiments on a well-known blog emotion corpus show that our proposed multi-label emotion intensity analysis algorithm outperforms baseline methods by a large margin.

Keywords: Opinion mining · Fuzzy relation equation · Sentiment analysis · Multi-labeled emotion · Emotion intensity · Fuzzy-rough set

1 Introduction

With the development of Web 2.0 techniques, more and more people are willing to express their feelings and emotions via the social media platform such as blog, microblog, and online forum. Therefore, detecting and analyzing the sentiments embedded in social media has become a popular research topic for both academic communities and commercial companies.

H. Gao et al. (Eds.): DASFAA 2016 Workshops, LNCS 9645, pp. 65–80, 2016.
DOI: 10.1007/978-3-319-32055-7_6

For sentiment analysis, previous researches usually focused on sentiment orientation classification, i.e., classifying the subjective text into two-orientation (*positive* and *negative*) or three-orientation (*positive*, *neutral*, and *negative*). However, currently more and more researches consider the sentiment categories such as *joy*, *hate*, *love*, *sorrow*, *anxiety*, *surprise*, *anger*, *expect* [11] called as fine-grained emotion (we call this research as emotion analysis in this paper). Compared with sentiment orientation classification, the fine-grained emotion analysis could capture users' meticulous sentiments (i.e., emotions), and are more suitable for public opinion monitoring about online hot events.

In fact, in social media platform, multiple emotions may be coexisting in just one document, paragraph, or even sentence, as shown in the following example.

> *"As a teacher, I'm very happy. I love my work, and like my school. However, the air quality of the city is too bad, and housing prices are so expensive."*

In above example, for sentiment orientation analysis task, the paragraph contains positive (expressed by sentiment word "love", "like", and "happy") and negative (expressed by sentiment word "bad" and "expensive") sentiments. For fine-grained emotion analysis, the paragraph simultaneously contains joy (expressed by sentiment word "happy"), love (expressed by sentiment word "love" and "like"), anger (expressed by sentiment word "bad"), and anxiety (expressed by sentiment word "expensive") emotions. For such analysis task, a multi-label fine-grained emotion analysis will be required. Moreover, for the same emotion love, the emotion intensity of word "love" and "like" is different.

Generally, sentiment orientation and emotion are expressed implicitly by sentence structure, semantic, and sentiment words including adjectives, verbs, and adverbs. Table 1 shows some examples. Here each post is associated with several different emotion labels and a value between $0 \sim 1$ indicates the intensity of every emotion. In Table 1, the third post has the labels of *joy*, *hate*, *love*, and the forth post has the labels *hate*, *anger*. If an author expresses stronger emotion, the corresponding post will have a higher intensity. Because the word "fantastic" expresses a more intense emotion than the word "OK", so the intensities of *joy* and *love* emotions in first post are higher than those of second post. Analyzing the emotions in social text needs to not only recognize the multi-label co-existing emotions but also calculate their corresponding intensities.

The emotion analysis problem that simultaneously considering multi-label fine-grained emotions and their corresponding intensity is really rarely studied in the previous literature. To tackle this challenge, we regard the multiple emotion intensity detection in social text as an uncertain classification problem, and propose an approach that combining fuzzy relation equation with fuzzy-rough set for solving the problem. We first calculate the fuzzy emotion intensity (expressed as range) of every sentiment word by solving a fuzzy relation equation, then utilize an improved fuzzy-rough set method to predict emotion intensity for the social subjective text at sentence, paragraph, and document level.

In our method, human emotions have eight basic kinds of categories as defined in Quan's research [11], and each one with ten levels of intensity which is annotated between 0.1 and 1. Due to the intrinsic characteristic of human language, different sentiment words may express different emotion intensity. Even the same sentiment word may

Table 1. The examples of multiple emotions with different intensities

Social text	*Joy*	*Hate*	*Love*	*Sorrow*	*Anxiety*	*Surprise*	*Anger*	*Expect*
The movie tonight is fantastic, the dinner sucks!	0.5	0.2	0.5	0	0	0	0.7	0
The movie tonight is OK, but I still looking forward to Avengers 2!	0.2	0	0.2	0	0	0	0	0.5
I like this phone's screen, but it has a short battery life.	0.1	0.3	0.3	0	0	0	0	0
This is a totally rubbish phone! It cost me 500$ but was broke in one month!	0	0.6	0	0	0	0	0.8	0

have different emotion intensity in different context. To tackle these problems, firstly, we put all the sentiment words in a fuzzy relation equation and calculate each word's emotion value that is mapped into an intensity range with upper and lower bounds. Then we model the words by our improved fuzzy-rough set method which considering both the bounds of the intensity ranges and the importance of the words. Finally the intensities of the emotions are tuned by fuzzy modifiers to determine the overall emotions embedded in the social text. Experiment results using a well annotated blog emotion dataset show that our proposed algorithm significantly outperforms other baselines.

The rest of the paper is organized as follows: Sect. 2 introduces fuzzy relation equation and fuzzy-rough set. In Sect. 3, we propose the fuzzy relation equation and fuzzy-rough set based multi-label, fine-gained emotion intensity prediction method. In Sect. 4, we show our experiment setup and results. We survey the related work with sentiment and emotion analysis in Sect. 5. Finally we conclude the paper and give the future work in Sect. 6.

2 Fuzzy Relation Equation and Fuzzy-Rough Set

In this paper, we use improved fuzzy-rough set based on fuzzy relation equation for our multi-label and fine-gained emotion intensity computing. We introduce the fuzzy relation equation and fuzzy-rough set in this section.

2.1 Rough Set

Rough set theory was proposed by Pawlak in 1982 [10], and it is useful to deal with uncertainty analysis problems. In traditional Pawlak's rough set theory [10], the pair (U, R) is called as an approximation space, where U is a universe and R is an equivalence relation on U.

Suppose R is an indiscernibility relation on U, with respect to R, equivalence class of an element x in U could be defined an as follows:

$$[x]_R = \{y|(x, y) \in R\} \tag{1}$$

The quotient set of U by the relation R is denoted by U/R, and

$$U/R = \{X_1, X_2, \ldots, X_m\} \tag{2}$$

Let $[x]_R = \{y \in U|(x, y) \in R\}$ be the equivalence class of $x \in U$, the Pawlak approximation space is defined as follows.

Let X_i ($i = 1, 2, \ldots, m$) be an equivalence class of R, and the equivalence classes of R are called elementary sets. If given an arbitrary set $X \in R$, X may be characterized by a pair of upper and lower approximations defined as follows [10, 22]:

$$\bar{R}(X) = \{x \in U, [x]_R \subseteq X\} \tag{3}$$

$$\underline{R}(X) = \{x \in U,\ [x]_R \cap X \neq \phi\} \tag{4}$$

2.2 Fuzzy Set and Fuzzy Relation Equation

Fuzzy set theory was proposed by Zadeh in 1965 [20]. U is a finite and non-empty set, and is called as universe. In this paper, the universe U is considered to be finite. Fuzzy set A is a mapping from U into the unit interval [0, 1]:

$$\mu : U \to [0, 1] \tag{5}$$

where for each $x \in U$, $\mu_A(x)$ is called as the membership degree of x in A. The fuzzy power set is denoted by $F(U)$ [20] showing the set of all fuzzy sets in the universe U.

The fuzzy relation equation was proposed by Ernest et al. [5] which is an equation of the form $H \circ X = P$, where H and P are fuzzy sets, X is a fuzzy relation, "$\circ$" is fuzzy inner product, for $h_i \in H$, $x_i \in X$, $H \circ X = \vee(H \wedge X)$, where $\wedge$ means fuzzy intersection, it defined as $h_i \wedge x_i = \min(h_i, x_i)$, and $\vee$ means fuzzy union. It is defined as: $h_i \vee x_i = \max(h_i, x_i)$. And $H \circ X = P$ stands for the composition of H with X.

For a given Fuzzy matrix $H = [h_{m\times l}] \in \mu$, $P = [p_{n\times l}] \in \mu$, we calculate fuzzy matrix $X = [x_{n\times m}] \in \mu$ to meet the formula $H \circ X = P$ as:

$$\begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \cdot \begin{bmatrix} h_{11} & \cdots & h_{1l} \\ \vdots & \ddots & \vdots \\ h_{m1} & \cdots & h_{ml} \end{bmatrix} = \begin{bmatrix} p_{11} & \cdots & p_{1l} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nl} \end{bmatrix} \tag{6}$$

For most fuzzy relation equation, if it has solution, the solutions always have the following characteristics:

(1) The solution of the equation always appears in the form of set.
(2) A fuzzy relation equation always has more than one solution.

2.3 Fuzzy-Rough Set

Fuzzy-rough set was first proposed by Dubois and Prade in 1990 [3]. A fuzzy subset $R \in F(U \times W)$ is seemed as a fuzzy binary relation between U and W, and $R(x, y)$ is defined as the degree of relation between x and y, where $(x, y) \in U \times W$ [17, 18].

Definition 1: U and W are two finite and nonempty universes. Suppose that R is an arbitrary relation from U to W, the triple (U, W, R) is called a generalized fuzzy approximation space. For any set $A \in F(U)$, the upper and lower approximations of A, $\bar{R}(A)$ and $\underline{R}(A)$, with respect to the approximation space (U, W, R) are fuzzy sets of U whose membership functions, and for each $x \in U$, are defined respectively as:

$$\bar{R}(A) = \vee_{y \in W}[R(x,y) \wedge A(y)],\ x \in U \tag{7}$$

$$\underline{R}(A) = \vee_{y \in W}[1 - R(x,y) \vee A(y)],\ x \in U \tag{8}$$

The pair $(\bar{R}(A), \underline{R}(A))$ is referred to as a generalized fuzzy rough set, and R is referred to as upper and lower generalized fuzzy rough approximation operators.

3 Improving Fuzzy-Rough Set Based on Fuzzy Relation Equation for Multi-label Emotion Intensity Prediction

Since the emotion intensity between [0, 1] can be regarded as fuzzy degree, we consider to use fuzzy relation equation and fuzzy-rough set for emotion intensity analysis. In this paper, we propose an improved fuzzy-rough set based method that combined with fuzzy relation equation. The overall framework is shown as Fig. 1.

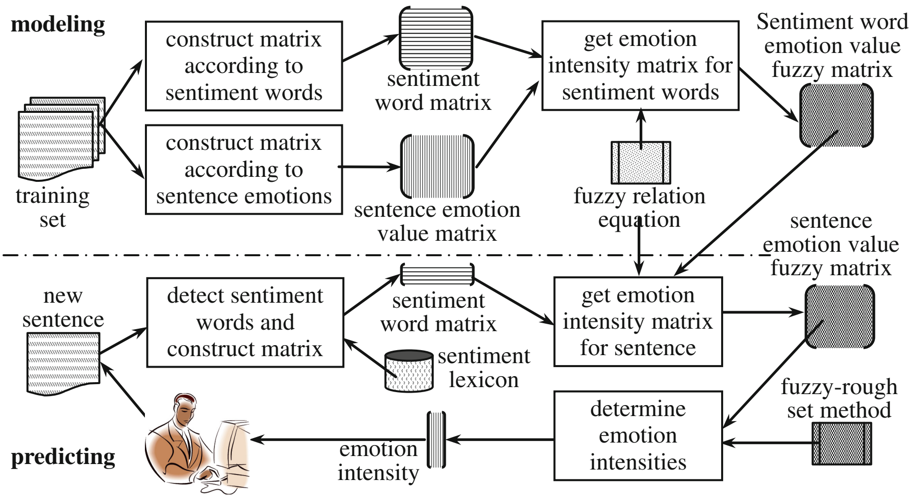

Fig. 1. Framework of multi-label emotion intensity analysis

In detail, in the modeling stage, we use an annotated blog set as training data in which all the sentiment words are labeled and the corresponding eight emotion labels and intensities are given for every sentence. Based on the training set, we can construct X matrix with sentiment words and P matrix with eight emotion intensities, and further calculate matrix H according to Formula (6) $H \circ X = P$. Here X can be regarded as eight emotion intensities of every sentiment word. As mentioned in Sect. 2.2, H is a fuzzy matrix. After modeling (obtaining matrix H), in the predicting stage, for a new sentence we can construct X by detecting sentiment word of the sentence, we further apply an improved fuzzy-rough set method for obtaining eight emotion intensities of the sentence. Moreover, the process can be extended to paragraph and document level.

The main techniques used in modeling and predicting stage include fuzzy relation equation and fuzzy-rough set (here we improve it) method, so in this section, we introduce their applications in our multi-label emotion intensity analysis.

3.1 Fuzzy Relation Equation Calculation

As mentioned above, our training set has been labeled emotion intensity at sentence level, and the embedded sentiment words are also known (even unknown, we can detect them with existing technique and sentiment lexicon). So in modeling stage, we can use a fuzzy relation equation (Formula (6)) to get emotion intensity fuzzy matrix of all sentiment words in the training set.

For clearer significance, here we rewrite fuzzy relation equation in Formula (6) and present it as Formula (9).

$$\begin{bmatrix} vw_{11} & \cdots & vw_{1n} \\ \vdots & \ddots & \vdots \\ vw_{m1} & \cdots & vw_{mn} \end{bmatrix} \cdot \begin{bmatrix} ve_{11} & \cdots & ve_{18} \\ \vdots & \ddots & \vdots \\ ve_{n1} & \cdots & ve_{n8} \end{bmatrix} = \begin{bmatrix} vs_{11} & \cdots & vs_{18} \\ \vdots & \ddots & \vdots \\ vs_{m1} & \cdots & vs_{m8} \end{bmatrix} \tag{9}$$

In above equation, the first item in left $VW = [vw_{mn}]$ is sentiment word matrix, here we assuming the training set T has n sentiment words and m sentences. If jth sentiment word w_j exists in ith sentence s_i, $vw_{ij} = 1$ else $vw_{ij} = 0$.

The item in right $VS = [vs_{m8}]$ is sentence emotion intensity matrix, here we consider eight emotions as the same as [11], i.e., e_1 = *joy*, e_2 = *hate*, e_3 = *love*, e_4 = *sorrow*, e_5 = *anxiety*, e_6 = *surprise*, e_7 = *anger*, e_8 = *expect*. For the ith sentence s_i, vs_{i1}, vs_{i2}, …, vs_{i8} represent the emotion intensity value of *joy*, *hate*, *love*, *sorrow*, *anxiety*, *surprise*, *anger*, *expect*, respectively. We can construct the matrix based on known multi-label emotion intensity of every sentence in T.

The second item in left $VE = [ve_{n8}]$ is emotion intensity matrix of all sentiment words. For the ith sentiment word w_i, ve_{i1}, ve_{i2}, …, ve_{i8} represent the corresponding eight emotion intensity values of w_i, respectively. In the modeling stage, our goal is just to calculate it by solving the fuzzy relation equation in Formula (9).

The following Algorithm 1 describes the process for achieving *VE* matrix.

Algorithm 1: Modeling Algorithm for Multi-label Emotion Intensity

---

Input: Training Set *T* // In *T*, all sentiment words have been labeled and eight emotion intensities of every sentence are known
Output: emotion intensity matrix *VE* about all sentiment words
Description:

1. For every sentence $s \in T$;
2. {Construct a row of *VW* matrix with all sentiment words in *s*;
3. Construct a row of *VS* matrix with eight emotion intensity values of *s*;
4. }
5. Solve and Return fuzzy matrix *VE* with fuzzy relation equation in Formula (9);

3.2 Improving Fuzzy-Rough Set

According to the framework in Fig. 1, in the predicting stage, for a new sentence *s*, we apply *VE* matrix returned by Algorithm 1 to achieve a new emotion intensity matrix *VE*'. Based on *VE*' we further calculate eight emotion intensity values corresponding to *s*. Because *VE* is a fuzzy matrix, *VE*' is also fuzzy. In this case, we further process it with an improved fuzzy-rough set for determining emotion intensity values for *s*. Here we introduce the improved fuzzy-rough set method and its application in our work.

Definition 2: (F^{-1}, W) is considered as a relation of a fuzzy set over universe E iff F^{-1} is a mapping of U into the set of all fuzzy subsets from the set W, where F^{-1} is a mapping given by F^{-1}: $E \rightarrow F(W)$. $F(E)$ denotes W, which is regarded as all fuzzy subsets of parameter set, then $F^{-1}(e)(w) \in [0, 1]$, $\forall e \in E$, $w \in W$ [8].

In this section, we will establish an improved model based on fuzzy-rough sets, which is able to calculate the weight of each attribute, i.e., the emotion intensity, for the given sentence. In Sect. 3.1, we associate each sentiment word with the eight basic emotions: *joy*, *hate*, *love*, *sorrow*, *anxiety*, *surprise*, *anger* and *expect*, and most sentences have more than one emotion words in this situation. We need a new algorithm to estimate the multi-label emotion intensity value of the whole sentence. The process can also extend to paragraph or document.

Definition 3: Let (F^{-1}, W) be a fuzzy set over E, the triple relation (E, W, F^{-1}) is called as the fuzzy approximation space. For any $A \in F(W)$, the upper and lower approximations of A, $\bar{F}(A)$ and $\underline{F}(A)$ with respect to the fuzzy approximation space (E, W, F^{-1}) are fuzzy sets of U whose membership functions, are defined as followings:

$$\bar{F}(A)(x) = \vee y \in W[(F^{-1}(x)(y)) \wedge A(y)],\ x \in E \tag{10}$$

$$\underline{F}(A)(x) = \wedge y \in W[(1 - F^{-1}(x)(y)) \vee A(y)], x \in E \tag{11}$$

After we find out attributes of sentiment words existing in sentence s, to make this algorithm more suitable for sentiment analysis task and human language logic, according to Definitions 1 and 4, we propose an improved fuzzy-rough set method. For any $A \in F(W)$, the upper approximation and lower approximation of A, $\bar{F}(A)$ and $\underline{F}(A)$. Under this specific situation of this paper, we defined $A = \{\bar{A}, \underline{A}\}$, because of the fuzzy relation equation the solution of VE in Formula (9) may be a range, which is described as $[F(e)(\underline{w}), F(e)(\bar{w})]$, then we defined $\bar{A}$ for $\bar{F}(A)$, $\underline{A}$ for $\underline{F}(A)$, and if $F(e)(w) = 0$ then $\bar{F}(A) = 0$ and $\underline{F}(A) = 0$, if $F(e) \neq 0$ then

$$\bar{F}(A)(e) = \vee_{w \in W}[F(e)(\bar{w}) \wedge \bar{A}(w)], e \in E \tag{12}$$

$$\underline{F}(A)(e) = \begin{cases} \wedge_{w \in W}[(1 - F(e)(\underline{w})) \vee \underline{A}(w)], & e \in E, F(e)(\underline{w}) \in [0.5, 1] \\ \wedge_{w \in W}[F(e)(\underline{w}) \wedge \underline{A}(w)], & e \in E, F(e)(\underline{w}) \in [0, 0.5) \end{cases} \tag{13}$$

In this paper, all the emotional intensities are between 0 and 1, so it is suitable for fuzzy degree. Here A is very important, and we choose the strongest emotion intensity of the solutions, which the upper bound is showed as $\bar{A}$, and lower bound as $\underline{A}$. Specifically, we will calculate the strongest emotion intensity which is belong to upper bound as the result of $\bar{A}$, and lower bound as the result of $\underline{A}$. the logic of the improving fuzzy rough algorithm is the more approximate to A, the stronger of the emotion intensity of the sentiment words is. So we can construct the decision object A on the evaluation of the words universe W.

Take one sentence for example in our training set: "我想, 人其实内心都有顽强的意志力的, 只不过有些人没释放出来而已 (*I think actually there is strong willpower in people" heart, but some people just have not release yet*)". We calculate $\bar{F}(A)$ and $\underline{F}(A)$ with Formula (12) and (13), the result is shown as Table 2.

Table 2. The examples of emotion computation based on our improved fuzzy-rough set

	顽强(strong)	意志力 (willpower)	只不过 (just)	而已 (yet)	$\bar{F}(A)$	$\underline{F}(A)$	intensity
Joy	0	0	0	0	0	0	0
Hate	0	0	0	0	0	0	0
Love	[0.4,0.7]	0.5	0	0	0.7	0.4	1.1
Sorrow	0	0	0	0	0	0	0
Anxiety	0	0	[0.4,0.7]	[0.2,0.5]	0.7	0.2	0.9
Surprise	0	0	0	0	0	0	0
Anger	0	0	0	0	0	0	0
Expect	0	0	[0.3,0.6]	[0.3,0.8]	0.8	0.3	1.1
$\bar{A}$	0.7	0.5	0.7	0.8			
$\underline{A}$	0.4	0.5	0.4	0.3			

In Table 2, four words are selected from the sentences by the sentiment lexicon, and the shadow part is the new emotion intensity value matrix, which is calculated from returned result *VE* by Algorithm 1 according to sentiment words in the sentence. In the dataset, the emotion labels and intensity values of the sentiment words are marked by people with 顽强 strong (*love* = 0.7), 意志力 willpower (*love* = 0.5), 只不过 just (*anxiety* = 0.4, *expect* = 0.3), and 而已 yet (*anxiety* = 0.3, *expect* = 0.3). Obviously, most these human annotated values are in the range of our predicted results. It can prove the effectiveness of our Algorithm 1.

Moreover, in the dataset, this sentence is annotated by the emotional intensity: *love* = 0.5, *anxiety* = 0.4 and *expect* = 0.5. By further normalization process with linear regression, our calculated values are the same order as the result annotated by human.

In summary, for a new sentence, the process of predicting its multi-label emotion intensity is shown as Algorithm 2.

Algorithm 2: Predicting Multi-label Emotion Intensity for a New Sentence

Input: a new sentence *s*, emotion intensity matrix *VE* from Algorithm 1
Output: emotion intensity values of *s*
Description:

1. Find all sentiment word $w \in s$;
2. Construct sentiment word matrix VW' with all above $w \in s$;
3. For every $w \in VW'$
4. Achieve emotion intensity value of w from VE';
5. Construct emotion intensity value fuzzy matrix W;
6. Compute strongest emotional intensity object A as:

$$\overline{A} = \sum_{i=1}^{W} \frac{maxF(\overline{w}_i)}{w_i}, w_i \in W, i.e., \overline{A}(w_i) = max\{F(e_j)(\overline{w}_i) \big| e_j \in E\}$$

$$\underline{A} = \sum_{i=1}^{W} \frac{maxF(\underline{w}_i)}{w_i}, w_i \in W, i.e., \underline{A}(w_i) = max\{F(e_j)(\underline{w}_i) \big| e_j \in E\}$$

7. Calculate the improving fuzzy-rough upper approximation $\overline{F}(A)$ and fuzzy-rough lower approximation $\underline{F}(A)$ // see formula (12) and (13);
8. Calculate the choice value $m = \overline{F}(A)(e_i) + \underline{F}(A)(e_i)$, $e_i \in E$;
   // E is the universe of the emotion
9. Normalize m into [0, 1] with Linear Regression Algorithm;
10. Return m as emotion intensity;

4 Experiments

For proving the validity and advantage, in this section we give comparison experiments.

4.1 Dataset and Evaluation Metric

In this section we use Quan's [11] Chinese blog dataset to evaluate our proposed method. The corpus contains 1,487 documents, with 11,953 paragraphs, 38,051 sentences, and

971,628 Chinese words. All sentiment words are labeled, and every sentence and sentiment word are annotated with eight basic kinds of emotions with intensities. An example is shown as Fig. 2 (which is actually the sentence of Table 2).

Although Quan's annotated dataset contains emotion intensity labels for each key sentiment word in blogs, our proposed fuzzy-rough set based method does not rely on the labels at the word level to predict the emotions in sentences, paragraphs and documents. In the training stage, all we need are the emotion intensity labels for the sentences in the training dataset and a sentiment lexicon. Therefore, we ignore the emotion labels of the sentiment words and their corresponding emotion intensities. We use 5 fold cross validation for the experiments. The dataset is divided into 5 parts on average, and each time there are four training sets and one test set.


```
<sentence S="我想，人其实内心都有顽强的意志力的，只不过有些人没释放出来而已。">
  <S_no>第7段第3句标注</S_no>
  <Segmented_S> 我想/n ，/w 人/n 其实/d 内心/n 都/d 有/v 顽强/a 的/u 意志力/n 的/u ，/w 只不过/d 有些/r 人/n 没/d 释放/v 出来/v 而已/y 。</Segmented_S>
  <S_Length>32</S_Length>
  <Keywords Anger="0" Anxiety="0" Expect="0" Hate="0" Joy="0" Love="0.7" Opinionholder="0" POS="a" Sorrow="0" Surprise="0" end="-1" position="10" start="-1">顽强
    </Keywords>
  <Keywords Anger="0" Anxiety="0" Expect="0" Hate="0" Joy="0" Love="0.5" Opinionholder="0" POS="n" Sorrow="0" Surprise="0" end="-1" position="13" start="-1">意志力
    </Keywords>
  <Keywords Anger="0" Anxiety="0.4" Expect="0.3" Hate="0" Joy="0" Love="0" Opinionholder="0" POS="d" Sorrow="0.0" Surprise="0" end="-1" position="18" start="-1">只不
    过</Keywords>
  <Keywords Anger="0" Anxiety="0.3" Expect="0.3" Hate="0" Joy="0" Love="0" Opinionholder="0" POS="y" Sorrow="0" Surprise="0" end="-1" position="29" start="-1">而已
    </Keywords>
  <noword end="-1" position="24" start="-1" modifier_word_length="2" modifier_word_position="25">没</noword>
  <Opinion_Fact>opinion</Opinion_Fact>
  <Polarity>积极</Polarity>
  <Opinion_holder end="0" position="0" start="0">我想</Opinion_holder>
  <Joy>0.0</Joy>
  <Hate>0.0</Hate>
  <Love>0.5</Love>
  <Sorrow>0.0</Sorrow>
  <Anxiety>0.4</Anxiety>
  <Surprise>0.0</Surprise>
  <Anger>0.0</Anger>
  <Expect>0.5</Expect>
```


Fig. 2. An annotated blog fragment in the training set

Since not all the traditional multi-label learning metrics could meet the need of the multi-label intensity prediction problem, in this experiment, we use four evaluation metrics [21], and some are revised versions of the formulas to fit our problem.

$$\textbf{Subset Accuracy}: subsetacc_s(h) = \frac{1}{p}\sum_{i=1}^{p}[|h(x_i = Y_i)|] \tag{14}$$

The subset accuracy evaluates the fraction of correctly classified examples, i.e. the predicted label set is identical to the ground-truth label set. Intuitively, subset accuracy can be regarded as a multi-label counter part of the traditional accuracy metric, and tends to be overly strict especially when the size of label space is large.

$$\textbf{Hamming Loss}: hloss_s(h) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{q}|h(x_i)\Delta Y_i| \tag{15}$$

where Δ stands for the symmetric difference between two sets. The hamming loss evaluates the fraction of misclassified instance-label pairs, i.e., a relevant label is missed or an irrelevant is predicted. This formula is suitable for the multi-label learning problem. In this paper, because we focus on multi-label intensities of multiple emotions, we need to revise the hamming loss metric. In the formula below, we try to measure the difference between the value we predict and the ground truth one:

$$hloss_s(h) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{q}|h(x_i) - \Delta Y_i| \tag{16}$$

$$\textbf{One-error}: \; One\text{-}error_s(h) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{q}[[\arg\max_{y\in Y} f(x_i, y)] \notin Y_i] \tag{17}$$

The one-error evaluates the fraction of examples whose top-ranked label is not in the relevant label set. In our paper, it means that the strongest emotion of the eight we predict is wrong. As we can see, the value of the One-error is the fewer, the better.

Average precision :

$$avgprec_s(h) = \frac{1}{p}\sum_{i=1}^{p}\frac{1}{Y_i}\sum\nolimits_{y\in Y_i}\frac{|\{y'|rank_f(x, y') \le rank_f(x_i, y), y' \in Y_i\}|}{rank_f(x_i, y)} \tag{18}$$

The average precision evaluates the average fraction of relevant labels ranked higher than a particular label $y \in Y_i$. In this paper, it means whether the descending order of the value of each emotion is right. This metric is the larger, the better.

4.2 Experiment Setup

As few related researches were proposed for multi-label emotion intensity, we will compare our method with the following methods that can be divided into two categories.

(1) **Using Word Emotion Intensity Labels**. This kind of methods means that we leverage the word emotion labels and their corresponding intensity in the training blog dataset to predict the emotions in the test set. Note that this will reduce the difficulty of the learning problem. These methods include:

- Fuzzy Union (FU for short). It is defined as: $(A \cup B)(x) = \max(A(x), B(x))$ for all $x \in X$. Taking the value of joy from Fig. 2 for example: Love (Fuzzy union) = max(顽强(*love*), 意志力(*love*)) = max (0.7, 0.5) = 0.7.
- Naïve Bayes (NB for short). We assumed that every emotion is independent from each other. Taking the value of *joy* from Fig. 2 for example: Love (Naïve Bayes) = *P*(顽强| *love*) * *P*(意志力| *love*) * *P*(*love*) = (6/12) * (5/17) * (12/25) = 0.07058.
- Multi-label Prediction(ML for short). In this method, we have already known the emotion labels of sentiment words which are annotated by people in the corpus. So we aggregate the labels of the words to predict the binary emotion labels (0/1) of the testing text.
- Our proposed fuzzy-rough set based method (FRS for short). In this case, we do not need the fuzzy relation equation to predict the emotions of the word and we can directly use Formulas (12) and (13) to predict the emotion intensities in the test set [15].

(2) **Ignoring Word Emotion Intensity Labels**. This kind of methods means that we ignore the labels and intensities of the words in the training dataset and only utilize the labels at the sentence level to train classifiers. These methods include:

- Regression Analysis (RA for short). We assume two vector spaces: B and W, then constructing an equation: $Y = BW + C$. The object function Y is the emotional value result of the training set, W is the sentiment words of the training set, B is the coefficient need to be learned, and C is the adjustment factor. We built 8 RA models for 8 kinds of emotions.
- Our proposed fuzzy relation equation and fuzzy-rough set based method (FRE-FRS for short). We leverage Algorithm 2 to predict the emotion intensities in the test set.

4.3 Experiment Results and Discussion

In the first experiment, we evaluate the label prediction accuracy with Subset Accuracy, which depends on fuzzy relation equation. Using our method, the percent of emotions figured out in the article is showed in Table 3.

Next we compare our multi-label method with other baseline algorithms mentioned in Sect. 4.2. Our method ignores the emotions and their intensity of sentiment words when training because of its unrealistic in most datasets. In this experiment, we also give the comparison results using the emotion intensity labels of the sentiment words during the training stage. The comparison methods include ML method that only using word emotions and sentence level labels, and FU and NB methods that using both word emotions and intensities with sentence level labels. The evaluation results at three textual levels are shown in Table 4.

Table 3. Label prediction accuracy (ignoring word emotions and intensities)

	Subset accuracy (RA)	*Subset accuracy* (FRE-FRS)
Document	0.59385	**0.68279**
Paragraph	0.62934	**0.76349**
Sentence	0.78421	**0.87326**

Hamming Loss is a measure of the value of general accuracy. Hamming Loss is the fewer, the better. So as we can see, in both using and ignoring word emotions with intensities, our method is better than others.

Table 4. Emotion label intensity analysis of Hamming Loss

	Using labeled word emotions and intensities				Ignoring labeled word emotions and intensities	
	ML	FU	NB	FRS	RA	FRE-FRS
Document	0.09845	0.06853	0.21445	**0.03929**	0.28650	**0.16316**
Paragraph	0.10659	0.05117	0.15133	**0.02598**	0.23486	**0.13468**
Sentence	0.11179	0.03927	0.10320	**0.01767**	0.16494	**0.10050**

One-error is a method to measure whether the maximum value we predict is one of the final results, and the results we get are shown in Table 5.

Table 5. Label intensity analysis of one-error

	Using labeled word emotions and intensities			Ignoring labeled word emotions and intensities	
	FU	NB	FRS	RA	FRE-FRS
Document	0.78649	0.76689	**0.04764**	0.79764	**0.39281**
Paragraph	0.58395	0.58807	**0.06903**	0.69162	**0.46293**
Sentence	0.84421	0.37633	**0.07377**	0.66771	**0.52043**

Average precision is widely used in many areas. According to the measure of Average precision [21], the results of the experiments at three levels of text are showed in Table 6.

Table 6. Label intensity analysis of average precision

	Using labeled word emotions and intensities			Ignoring labeled word emotions and intensities	
	FU	NB	FRS	RA	FRE-FRS
Document	0.67249	0.36975	**0.74374**	0.24649	**0.40659**
Paragraph	0.71452	0.44310	**0.77940**	0.33980	**0.46698**
Sentence	0.70985	0.58617	**0.95494**	0.36796	**0.58838**

According to Formulas (17) and (18), one-error is the fewer, the better. In contrast, average precision is the larger, the better. So we can see our algorithm is significantly better than the other methods.

In Fig. 2, we see that all sentiment words have been labeled in the training dataset, and the emotion intensity values of both every sentiment word and every sentence are given. In this case, the evaluation results should be better than the ones without word emotion intensity labels. The results in Tables 4, 5, and 6 have validated our assumption.

However, it is difficult to get such kind of training set. In most cases, labeling emotion intensity for all sentiments words and sentences is unrealistic. Our method can model and predict emotion intensity for a new sentence and get better results without emotion labels and corresponding intensities in training set. Especially, when comparing with the method of regression analysis that ignoring word emotions and intensities, our proposed algorithm shows obvious advantages.

Generally, we argue that the fuzzy logic is suitable for multi-label emotion intensity analysis, which means it is consistent with the logic of human language when expressing emotions. In most related bibliographies with fuzzy mathematics, the introduced examples always depicted intensity analysis of human feeling, such as the oldness degree of 40 years old, or the height degree of a 180 cm man. These questions all got good solutions by using fuzzy logic. As we can see in the Table 2, the emotional logic is not a simple summation. A sentence may have one or more key emotion. In our method, we compare the values of every kind of emotions that may be upper bound or lower bound to the largest value of the word (shown as A) which it belong to. Because the value of sentiment words in different text may not be identical, we do not put them in isolation, but make them interact between each other, which is achieved by the advanced fuzzy rough set. The logic of our proposed algorithm is straightforward. The upper approximation we calculated indicates the most optimistic closeness to the largest emotion intensity, so we take upper bound into this consideration. Similarly, the lower approximation is the most conservative closeness. Finally we combined these two results together to solve the problem. Our model is more suitable for smaller text units, such as sentences and paragraphs. That is because when the text unit is larger, the embedded human emotions are more complex and confused.

5 Related Work

Our work is about multi-label emotion analysis and emotion intensity calculating. For the calculating, we apply fuzzy relation equation and fuzzy-rough set method.

The sentiment analysis researches can be dating back to the early of this century. Pang and Lee [9] showed the effectiveness of classification of emotion by using machine learning methods. Costa et al. [2] verified a method to combine mining algorithms and software agent to build blogs based on sentiment applications. Zhang et al. [23] proved a model to extract emotional characteristics from reviews of products as a weakness finder. Huang et al. [7] proposed a sentiment space model to deal with sentiment classification task by using the semantic information.

Multi-label emotion analysis can be seemed as multi-label learning problem. Fürnkranz et al. [6] proposed a problem transformation method which is called "Calibrated Label Ranking" could draw on the advantages of the pairwise preference learning and the conventional relevance classification technique, in which a separate classifier could be trained for distinguishing whether a label was relevant or not. Boutell et al. [1] built another common problem transformation algorithm called "Binary Relevance". It could transform the multi-label classification problem into the binary classification problem. Some scholars also proposed many other algorithms and methods in the multi-label learning area. This kind of methods makes the traditional supervised machine learning algorithms more suitable to deal with the multi-label problem. Elisseeff and Weston [4] improved the kernel learning algorithm SVM to solve the multi-label data problem.

Although there are a lot of research for multi-label learning and sentiment analysis, little work is done for the emotion intensities in one social media text post. In this paper, our work is regarding the problem as uncertain emotion classification and solving the problem.

Since the fuzzy-rough theory was proposed during the late of 20th century [3], it has been widely used in many areas, especially uncertain classification problems. Wang et al. [16] proposed a new uncertain classification algorithm which combined fuzzy-rough set with decision tree. Shi and Gong [12] built a model for uncertainty characterization called covering-based rough sets by using the advanced fuzzy-rough set. Xiao et al. [19] proposed a method to classify and predict whether the listed companies have financial distress based on the combination of fuzzy-rough set and D-S evidence theory. Sun and Ma [13] gave an approach to decision making problem by combining the soft set with fuzzy-rough theory.

Although fuzzy-rough theory has no existing literature on uncertain emotion classification for social media, some researchers such as Vincenzo and Sabrina [14] has already realized fuzzy logic is associated and consistent with human emotions.

6 Conclusion and Future Work

In this paper, we proposed a new way to solve the multi-label and fine-grained emotion intensity analysis problem. For this particular problem, we used a fuzzy relation equation and an improved fuzzy-rough set theory to model and predict emotion intensity of a sentence, paragraph, and document.

In the future, we would like to build the model of multi-label emotion and intensity analysis on microblog, website reviews, or other social media. The role of adverbs and negative words can be further taken into consideration, which can further improve the performance of multi-label emotion intensity analysis.

Acknowledgments. Project supported by National Natural Science Foundation of China (61370074, 61402091), the Fundamental Research Funds for the Central Universities of China under Grant N140404012.

References

1. Boutell, M., Luo, J., Shen, X., Christopher, M.: Brown: learning multi-label scene classification. Pattern Recogn. (PR) **37**(9), 1757–1771 (2004)
2. Costa, E., Ferreira, R., Brito, P., et al.: A framework for building web mining applications in the world of blogs: a case study in product sentiment analysis. Expert Syst. Appl. (ESWA) **39**(5), 4813–4834 (2012)
3. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. Int. J. Gen. Syst. **17**(2–3), 191–209 (1990)
4. Elisseeff, A., Weston, J.: A kernel method for multi-labelled classification. In: NIPS 2001, pp. 681–687
5. Ernest, C., Józef, D., Witold, P.: Fuzzy relation equations on a finite set. Fuzzy Sets Syst. **7**, 89–101 (1982)
6. Fürnkranz, J., Hüllermeier, E., Mencía, E., Brinker, K.: Multilabel classification via calibrated label ranking. Mach. Learn. (ML) **73**(2), 133–153 (2008)
7. Huang, S., You, J., Zhang, H., Zhou, W.: Sentiment analysis of Chinese micro-blog using semantic sentiment space model. In: Conference: Computer Science and Network Technology (ICCSNT), pp. 1443–1447 (2012)
8. Maji, P., Biswas, R., Roy, A.: Soft set theory. Comput. Math. Appl. **45**(4–5), 555–562 (2003)

9. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? Sentiment classification using machine learning techniques. CoRR cs.CL/0205070 (2002)
10. Pawlak, Z.: Rough sets. Int. J. Parallel Prog. (IJPP) **11**(5), 341–356 (1982)
11. Quan, C., Ren, F.: A blog emotion corpus for emotional expression analysis in Chinese. Comput. Speech Lang. (CSL) **24**(4), 726–749 (2010)
12. Shi, Z., Gong, Z.: The further investigation of covering-based rough sets: uncertainty characterization, similarity measure and generalized models. Inf. Sci. (ISCI) **180**(19), 3745–3763 (2010)
13. Sun, B., Ma, W.: Soft fuzzy rough sets and its application in decision making. Artif. Intell. Rev. (AIR) **41**(1), 67–80 (2014)
14. Vincenzo, L., Sabrina, S.: Knowl.-Based Syst. **58**, 75–85 (2014)
15. Wang, C., Feng, S., Wang, D., Zhang, Y.: Fuzzy-rough set based multi-labeled emotion intensity analysis for sentence, paragraph and document. In: Li, J., Ji, H., Zhao, D., Feng, Y. (eds.) NLPCC 2015. LNCS, vol. 9362, pp. 444–452. Springer, Heidelberg (2015). doi:10.1007/978-3-319-25207-0_41
16. Wang, X., Zhai, J., Lu, S.: Induction of multiple fuzzy decision trees based on rough set technique. Inf. Sci. (ISCI) **178**(16), 3188–3202 (2008)
17. Wu, W., Mi, J., Zhang, W.: Generalized fuzzy rough sets. Inf. Sci. (ISCI) **151**, 263–282 (2003)
18. Wu, W., Zhang, W.: Constructive and axiomatic approaches of fuzzy approximation operators. Inf. Sci. (ISCI) **159**(2), 233–254 (2004)
19. Xiao, Z., Yang, X., Pang, Y., Dang, X.: The prediction for listed companies' financial distress by using multiple prediction methods with rough set and Dempster-Shafer evidence theory. Knowl.-Based Syst. (KBS) **26**, 196–206 (2012)
20. Zadeh, L.: Fuzzy sets. Inf. Control (IANDC) **8**(3), 338–353 (1965)
21. Zhang, M., Zhou, Z.: A review on multi-label learning algorithms. IEEE Trans. Knowl. Data Eng. (TKDE) **26**(8), 1819–1837 (2014)
22. Zhang, W., Liang, Y., Xu, P.: Uncertainty Reasoning Based on Inclusion Degree. Tsinghua University Press, Beijing (2007)
23. Zhang, W., Xu, H., Wan, W.: Weakness finder: find product weakness from Chinese reviews by using aspects based sentiment analysis. Expert Syst. Appl. (ESWA) **39**(11), 10283–10291 (2012)