INTELLIGENT AUTOMATION & SOFT COMPUTING, 2017
http://dx.doi.org/10.1080/10798587.2016.1267243

Building ontology for different emotional contexts and multilingual environment in opinion mining

Wan Tao[a,b] and Tao Liu[a,b]

[a]School of Computer and Information Science, Anhui Polytechnic University, Wuhu, China; [b]Key Laboratory of Computer Application Technology, Anhui Polytechnic University, Wuhu, China

ABSTRACT
With the explosive growth of various social media applications, individuals and organizations are increasingly using their contents (e.g. reviews, forum discussions, blogs, micro-blogs, comments, and postings in social network sites) for decision-making. These contents are typical big data. Opinion mining or sentiment analysis focuses on how to extract emotional semantics from these big data to help users to get a better decision. That is not an easy task, because it faces many problems, such as different context may make the meaning of the same word change variously, at the same time multilingual environment restricts the full use of the analysis results. Ontology provides knowledge about specific domains that are understandable by both the computers and developers. Building ontology is mainly a useful first step in providing and formalizing the semantics of information representation. We proposed an ontology DEMLOnto based on six basic emotions to help users to share existed information. The ontology DEMLOnto would help in identifying the opinion features associated with the contextual environment, which may change along with applications. We built the ontology according to ontology engineering. It was developed on the platform Protégé by using OWL2.

KEYWORDS
Ontology; opinion mining; social media; emotional context; multilingual environment; OWL

1. Introduction

Today large amounts of information are supplied by social media users. They share information about products, services and their experiences on the social networks. The information may contain users' sentiments, such as joy, sad, like, or dislike (Sentiment and emotion mean similar, in the following, we usually use emotion in a more specific situation). User's sentiment orientation and emotion of the topic or event cannot only provide decision-making basis in business, but also provide support for government's public opinion monitoring. It is very hard to integrate users' comments and feedback into normal application system. Currently, opinion mining is increasingly important than ever before, especially in customer preference analysis and prediction. Most opinion mining attempts to identify the polarity of sentiment in three categories; positive, negative or neutral, but how to identify the polarity of sentiment is a difficult task. The same word in different contexts may convey different emotions. For example, the word "high" in a sentence "My salary is high" represents the positive emotion, but in "The price is high" may represent the negative emotion. Otherwise, multilingual applications have produced many problems, such as how to map a word in Chinese to a similar word in English.

In this paper, we construct an ontology DEMLOnto to help to solve the mentioned problems. This ontology would help in identifying the opinion features associated with the different contextual environments that contain positive or negative sentiments. Also some methods are proposed to deal with multilingual environment. The remaining portion of the paper is organized as follows: Section II, firstly introduces the special environment of social media, then describes sentiment and opinion mining, provides the background research used for why and how to integrate the ontology into opinion mining. Section III provides the definition of our ontology DEMLOnto, and focuses on how to design our ontology on emotions and multi-languages. Section IV provides constructs of the ontology in OWL2(a Web Ontology Language), focuses on discussion about how to design ontology, explains how to get new describing construct, provides the building process on the Protégé, and introduce how to use DEMLOnto to extract information from the social media, and Section V briefs on conclusion and future work.

2. Related work

2.1. Social media environment

Social media supports the interaction among people in which they create, share, discuss, or exchange ideas in virtual communities and networks. This kind of online interaction has formed a culture of its own and has altered interpersonal communication of individuals, communities and societies all over the world. e.g. Sentiment expression in micro-blog posts often reflects user's specific individuality due to different language habits, personal character, opinion bias and so on (Wu, Song, & Huang, 2015).

Social media depends heavily on mobile and web-based techniques to create highly interactive platforms. It brings us with typical big data. Firstly, it produces large amounts of data. For instance, daily online QQ users count for 160 million. Their storage is 300G. Sina's daily traffic is 1 billion, and peak response is 1 million per second. Secondly, data update

CONTACT Wan Tao taowan@ahpu.edu.cn

fast; Sina's micro blog posts about 2500 more per second and Twitter posts about 14300 per second. All these data contain dynamic features and evolution characteristics. Thirdly, social media data is extremely various. This kind of variety is not only limited into its data type, which is structured or mainly unstructured, but also means the following: Many new words produced by media users (e.g. "童鞋" means "同学", which means "classmate" in English); kinds of contents inserted by supplying simple "Like" or "Dislike" tools, star-rating systems, tag-based annotation and navigation, and so forth. Lastly, there is uncertainty in social media data. Social media users exhibit various styles when expressing their feelings online. It is hypothesized that such diversity of sentimental manifestation may be pertinent to the latent aspects of different people including their personality, educational background, current mood and some unknown factors. While using the same word, people may deliver different sentiment orientations depending on the underlying context (Song, Feng, & Gao, 2015), just as our example in Section 1. Also for another example, the sentiment word "long" in the context of movies represents a negative sentiment when the movie is boring, whereas in the context of computer games maybe depict positive sentiment. Our work mainly has relationship with the last two points. Now, we especially focused on the uncertainty brought by the vocabulary.

From the above analysis, we found that vocabulary is the basis of sentiment analysis. When we do sentiment analysis, first of all, we should have a reserve of lexical knowledge, such as records of object features and corresponding subjective points of view, so that the massive sentiment analysis does not require relearning subjective points of view, then the efficiency of sentiment analysis is improved greatly. Ontology knowledge base is a good vocabulary knowledge base, which can help to understand concepts and the relationship among all the attributes in the concepts.

2.2. How to do opinion mining

Opinions are central to almost all human activities and are key influencers of our behaviors. Whenever we need to make a decision, we want to know others' opinions. In the real world, businesses and organizations always want to find consumer or public opinions about their products and services. Opinion mining, also called sentiment analysis, is the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. The task of determining positive and negative orientations of the information present in textual form is considered as a fundamental issue in opinion mining.

In general, opinion mining has been investigated mainly at three levels; document level, sentence level, entity and aspect level. It can be used for three varied objectives: Polarity identification, subjectivity or objectivity identification and feature/aspect based analysis (Dinakar, Andhale, & Rege, 2015). In this paper we focuses on opinions, which express or imply positive or negative sentiments, then make a summary at different levels.

Opinion mining consists of the following tasks (Liu & Zhang, 2012), p. (1) entity extraction and categorization, which is often done by ETL technology; (2) aspect extraction and categorization which mainly focus on opinion features; (3) opinion holder extraction and categorization; (4) time extraction and standardization; (5) aspect sentiment classification, which often determines the actual feeling about the mining object by assign a numeric sentiment rating to the aspect; (6) outputting the analysis results in tuples.

2.3. Opinion and emotion description

We can classify opinions as explicit opinion and implicit (or implied) opinion based on how they are expressed in text. Our research has focused on explicit opinions. According to Liu and Zhang (2012), an opinion can be defined as a quintuple

$$(e_i, a_{ij}, s_{ijkl}, h_k, t_l)$$

Where e_i is the name of an entity, a_{ij} is an aspect of e_i, s_{ijkl} is the emotional on aspect a_{ij} of entity e_i, h_k is the opinion holder, and t_l is the time when the opinion is expressed by h_k. The emotional s_{ijkl} is positive, negative, or neutral, or expressed with different strength/intensity levels, e.g. 1 to 5 stars as used by most review sits on the Web.

That is, the opinion s_{ijkl} must be given by opinion holder h_k about aspect a_{ij} of entity e_i at time t_l. Any mismatch is an error. At last, the feature pairs and emotional words can help us make decision.

Emotions are our subjective feelings and thoughts. Based on Parrott (2001), people have six primary emotions, i.e. love, joy, surprise, anger, sadness, and fear, which can be sub-divided into many secondary and tertiary emotions. Each emotion can also have different intensities. Emotions are closely related to sentiments. The strength of an emotion or opinion is typically linked to the intensity of certain emotions, e.g. joy and anger.

We think if we can integrate features (i.e. emotions, aspects, etc.) into the ontology, we will share them more comfortable. In Section 3 we will focus on emotional context, polarity strength and multi-language these aspects by integrating them into ontology.

2.4. Ontology description

Ontology has been defined as the specialization of the conceptualization by Gruber (1993). The main aim of ontology is to provide knowledge about specific domains that are understandable by both the computers and developers. Ontology plays important roles in sharing sources and defining terms precisely for future uses such as meta-data. It also helps to interpret a text review at a finger granularity with shared meanings and provides a sound semantic ground of machine understandable description of digital content.

The most common language to formalize ontology is OWL, a proposal of the W3C. Ontology based on Description Logics paradigm includes definitions of concepts (also as OWL classes), roles (also as OWL properties) and individuals. See details in Section 4.

In former research, Yaakub, Li, and Zhang (2013) proposed a feature ontology that uses a multidimensional model to integrate customer's characteristics and their comments about products. This approach first identifies the entities and then emotions present in the customers reviews related to mobiles are transformed into an attribute table by using a 7 point polarity system (−3 to 3). The limitation of their approach is that the ontology proposed by them is too general and is lack of reasoning ability among multiple products. According to Meersman (1999), most of the ontologies in the community of information systems are known as data models that are mainly used for structuring a fairly narrow application domain. It claimed that "ontology" that includes lexicons and thesauri may be a useful first step in providing and formalizing the semantics of information representation. In Thakor and Sasi (2015), a novel Ontology-based Emotion Analysis Process for Social Media content with negative emotions is presented. By using the OOP concepts, the noun and verb are used as object and object property respectively. The information on objects and object properties were used to build the ontology model. In Luo, Xu, Yu, and Chen (2011), Hu, Xu, Liu, and Mei (2014), Xu et al. (2015), Xu et al. (2016a, 2016b, 2016c), a whole framework for building domain ontology of video is proposed. The semantic link network model is used to mine and organize video resources based on their associations. A semantic-based video organizing platform is provided for searching videos.

However, we found that even previous researches have focused emotion or semantic analysis at different levels but domain knowledge, context relationship, and multilingual application were not considered during those researches. We think if ontology can be integrated into the 6 tasks of opinion mining (introduced in Section 2), it will make the mining results to be structured, make the results easy to be utilized and shared. Here, Figure 1 introduces how we integrate the ontology building into the general opinion mining process.

3. The ontology of emotions and multi-language

3.1. How to deal with emotions

We have developed DEMLOnto ontology ascended from Francisco, Gervás, and Peinado (2007) and Baldoni, Baroglio, and Rena (2012) with some changes. Our ontology is designed on six basic emotions: Love, joy, anger, sad, fear, surprise. They are structured in a taxonomy that covers from six basic emotions concepts to the most specific emotional words as data objects or instances. Each of the emotional data is related with the seven emotional strength levels by means of data ranges, as well as three polarity values to stand for positive, negative, or neutral.

Definition 1: Ontology is defined as:

$$O = \{C, A^C, R, A^R, H, I, X\}$$

where C is a set of concepts; A^C is a set of multiply properties, which one is belongs to a concept; R is a set of relationships; A^R is a set of multiply relationships, which one is belongs to a property; H is a set of hierarchy relationships among concepts; I is a set of instances; X is a set of axioms.

In Figure 2, we explain how concept Joy and Sad are constructed into the emotional ontology. In particular, when the application environment is not the same, we will make record of the different emotion-feature pairs into the ontology base, such as (price, i-Phone, highness, -1), (Tom's salary, incoming, highness,+2).

Figure 3 is a fragment of our ontology DEMLOnto's definition. In the definition, we use many properties to explain the feature of an emotion. For instance, polarity of an emotion, which value is [1, 0, -1] respectively, represents the emotion is positive, negative, or neutral; power represents the degree of positive or negative. We use (-3, 3), there are totally 7 numbers to stand for the different degrees of an emotion Why do we design two properties to measure an emotion? We do it just for that different applications maybe require different granularity in analysis process.

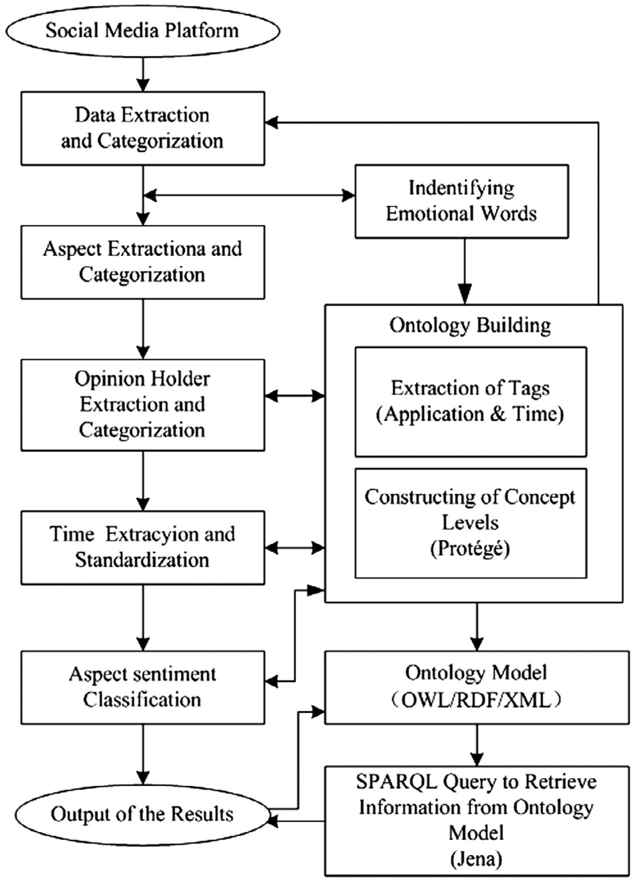

Figure 1. Opining mining with ontology building.

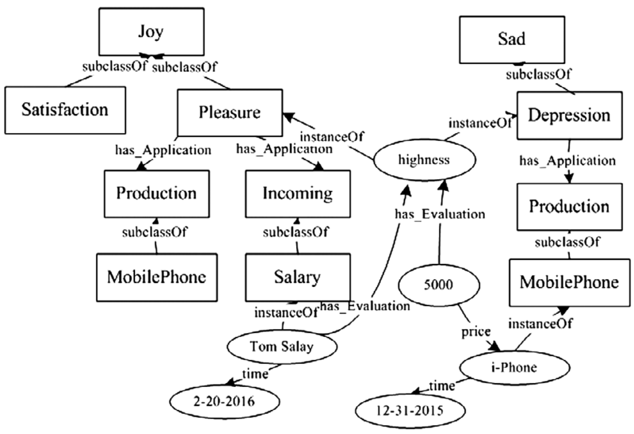

Figure 2. Emotional ontology (fragment).

O_{em}= {C_{em}, $A_{em}{}^{C}$, R_{em}, $A_{em}{}^{R}$, H_{em} ,$I_{em}{}^{C}$,X_{em} } *where*
C_{em}= { *Joy, Anger, Sad, Fear, Surprise, Love* }
$A_{em}{}^{C}$={ $A_{em}{}^{C}$ *(Joy)*, $A_{em}{}^{C}$ *(Anger)*,…}
$A_{em}{}^{C}$ *(Joy)*={ *polarity(integer),power(integer)*}
$R_{em}{}^{C}$ *(Joy)* ={ *subclassOf, instanceOf ,is_Similarwith, has_Application, has_Apptime, has_Evaluation….*,}
$A_{em}{}^{R}$ *(Joy,is_Similarwith)={inChinese(text), inEnglish(text), inSpanish(text)…}*
H_{em}*=(Is-a, Superclass, Subclass…)*
$I_{em}{}^{pleasure}$*={highness, goodness, fineness…}*
X_{em}*={polarity_rules, power_rules…}*
…

Figure 3. Ontology DEMLOnto definition (fragment).

3.2. How to deal with multi-languages

In order to deal with multiple languages in some application, we extended the original ontology with some changes from Baldoni, Baroglio, and Rena (2012). The ontology has three root classes: *Emotion*, *Word, and Application*. *Emotion* is the root for all the emotional concepts. *Word* is the root for the emotion-denoting words, i.e. the words which each language provides for denoting emotions, easily to be extended. Now we supposed it originally had two subclasses: *EnglishWord* and *ChineseWord*, All original words can get by information extraction from the Web (just as task 1 of opinion mining). *Application* includes objects, which the emotions come from.

And in Figure 3 we designed *is_Similarwith* relationship to help us dynamically classify the new words, which are provided for denoting emotions into *Word* concept. The ontology's structure will be listed in Figure 5, which is worked out by Protégé in Section 4.

4. Building the ontology in OWL2

4.1. OWL2 overview

OWL[1] is a recommendation of the W3C from 2004. In 2009, OWL 2 (W3C Recommendation, 2009) is proposed as an extension and revision of the OWL. OWL2 also has three increasingly-expressive sublanguages: OWL Lite, OWL DL, and OWL Full. Just as the former OWL. OWL Lite is simple, but a lack of logical expression ability; OWL Full is powerful for representing complex statements, but not useful for reasoning with them due to their computational properties. OWL DL is the subset of OWL designed for applications that need the maximum expressiveness without losing computational completeness and decidability. It is based on Description Logics, a particular fragment of first order logic, in which concepts, roles, individuals and axioms that relate them are defined.

In our work, we apply OWL DL to construct our ontology. In Table 1 we listed some constructs we have used, some of them are little different from the former OWL edition.

4.2. Building process of DEMLOnto

According to Noy and Mcguinness (2001), building an ontology includes: (1) defining classes in the ontology; (2) arranging the classes in a taxonomic (subclass–superclass) hierarchy; (3) defining properties and describing allowed values for these properties; (4) filling in the values for properties for instances. A little difference from this standard one, our building process has some interaction with opining mining, just as Figure 1 has shown.

At first, English emotional classes were extracted as the opinion as positive and negative based on the scales developed in previous works by Yaakub, Li, and Zhang (2013), and Chinese classes were gained from Xu, Lin, and Mei (2008). All of these are read and evaluated by human beings.

We use Protégé 5.0[2] as the building platform. Protégé is mainly an ontology editor developed by Stanford University and it has many plug-in reasoners, which can automatically classify the concepts into a hierarchy of emotional concepts.

In OWL2, we can define new constructs and axioms to fit for our needs. For instance, in Section 3 we have designed *is_Simliarwith* to handle similar words in different language contexts. We use *subClassOf, Domain* and *Range, etc.* to define *is_Simliarwith* just as an *ObjectProperty* of class *Emotion*. It

Table 1. OWL 2 constructs.

Functional Syntax	RDF Syntax
SubClassOf(C_1 C_2)	C_1 rdfs:subClassOf C_2.
EquivalentClasses(C_1 C_n)	C_j owl:equivalentClass C_{j+1}. j = 1…n-1
DisjointClasses($C_1$$C_2$)	C_1 owl:disjointWith C_2.
DisjointUnionOf(CN C_1… C_n)	CN owl:disjointUnionOf (C_1… C_n).
SubObjectPropertyOf(P1 P2)	P1 rdfs:subPropertyOf P2
EquivalentObjectProperties(P1 … Pn)	Pj owl:equivalentProperty Pj+1. j = 1…n-1
DisjointObjectProperties(P1 P2)	P1 owl:propertyDisjointWith P2
InverseObjectProperties(P1 P2)	P1 owl:inverseOf P2
SubDataPropertyOf(R1 R2)	R1 rdfs:subPropertyOf R2
DisjointDataProperties(R1 R2)	R1 owl:propertyDisjointWith R2
ObjectPropertyDomain(P C)	P rdfs:domain C
ObjectPropertyRange(P C)	P rdfs:range C
DataMaxCardinality(n R)	_:x rdf:type owl:Restriction. _:x owl:onProperty R _:x owl:maxCardinality n
DataMinCardinality(n R)	_:x rdf:type owl:Restriction. _:x owl:onProperty R _:x owl:minCardinality n

Note: CN represents class name; _:x represents blank node.

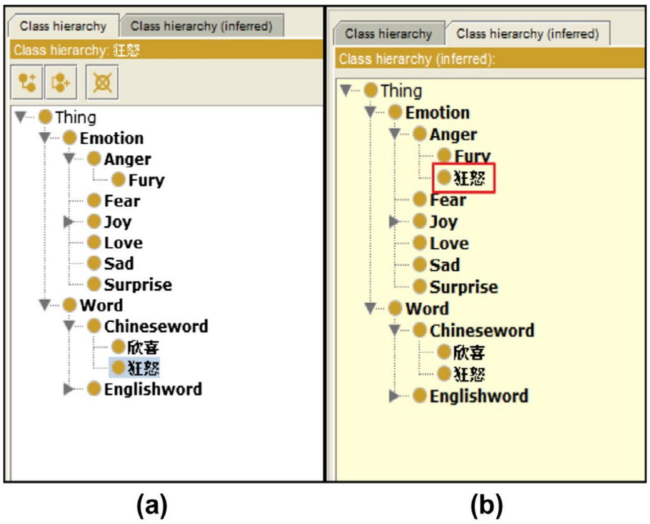

Figure 4. Reasoning in multi-language.

Table 2. ObjectProperty of DEMLOnto (Part).

Property Name	Class Name	Domain	Range
has_Application	*Emotion*	*Emotion*	*Application*
has_Apptime	*Application*	*Emotion*	*Datatime*
has_Evulation	*Application*	*Emotion*	*Integer*
is_Similarwith	*Emotion*	*Emotion*	*Word*

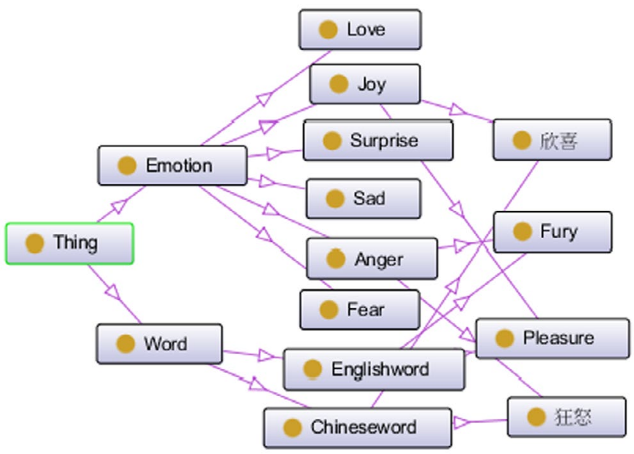

Figure 5. DEMLOnto structure (fragment).

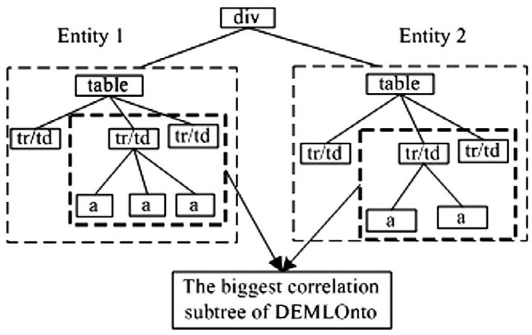

Figure 6. Extraction of similar entities.

helps us map the Chinese words to similar English words, and then they can be classified into the correct emotional class by the reasoner. The results are shown in Figure 4. We can see in the Figure 4(a), "狂怒" is a subclass of *Chineseword* ("狂怒 means "flying into rage", which has the similar meaning with "fury"). At the same time "狂怒" has property *is_Simliarwith* with *"Fury"* that is a *subclass* of *Anger.* As seen in Figure 4(b), "狂怒" is automatically added into subclasses of *Anger by* the reasoner Pellet (a plug-in of Protégé). Otherwise, we also defined *has_application, has_apptime, has_evaluation, etc.* properties in *ObjectProperty* of different concepts. We show them in Table 2.

The ontology structure is shown in Figure 5, which is displayed by OntoGraf (a plug-in of Protégé). In Figure 5, Chineseword and Englishword are subclassof Class Word, but Chineseword and Englishword are disjoint. In OWL2, such relationship can easily be represented in the following:

<!–*http://www.semanticweb.org/ DEMLOntology#Chineseword* –>
<owl:Class rdf:about="http://www.semanticweb.org/ *DEMLOntology #Chineseword">*
<rdfs:subClassOf rdf:resource="http://www.semanticweb.org/ *DEMLOntology #Word"/>*
<owl:disjointWith rdf:resource="http://www.semanticweb.org/ *DEMLOntology #Englishword"/>*
</owl:Class>

In summary, the ontology *DEMLOnto* can be supplementing accumulatively under the help of opinion extraction. Conversely, the built ontology can fix the mining results into common knowledge to help us eliminate ambiguity in opinion mining.

Table 3. The comparison of different methods.

Method	Domain	P(%)	R(%)	F (%)
Our method	Book	94.2	93.8	94.0
	Weather	93.3	92.0	92.6
	Shopping	91.5	90.8	91.1
DOM tree only	Book	91.4	90.1	90.8
	Weather	89.3	88.9	89.1
	Shopping	86.5	85.8	86.2

4.3. How to use DEMLOnto

As Figure 1 has shown, ontology can help us to extract effective opinion, to make us mine more useful results.

Now, we have used *DEMLOnto* to help us to extract opinion from micro-blog on domains of weather, book, and shopping. Under the guidance of the emotion ontology, we annotate entities at DOM level based on context distance and co-occurrence number. The annotation results were added to the entity record and similarity of the ontology and entity is calculated by using Formula 1. The extraction method and algorithm are shown in Figures 6 and 7.

$$Sim(C,O) = \frac{\max \sum_{j=1}^{n} \mathrm{Rel}(C, A_j)}{\sqrt{mn}} \tag{1}$$

Input: the ontology *DEMLOnto* and DOM tree of entities in the extracting web page
Output: the similar entities stored in T

1 $T \leftarrow R_i \in R$, $k \leftarrow 1$; //R is roots set of DOM tree, R_i is the ith root
2 repeat
  2.1 $T_j \leftarrow$The child node with the largest similarity as ontology;
  2.2 if Sim(C,O) is larger than threshold τ, and $Sim(T_j,O)>Sim(T,O)$
    then
      $T \leftarrow T_j$;
    else
      break;
  2.3 end if
3 until T_j is a leaf node;
4 repeat
  4.1 $T_l \leftarrow$ The leaf node of T;
  4.2 If Sim(C,O) is larger than threshold τ, and $|Sim(T+T_l,O) - Sim(T,O)| < \varepsilon$
    then
      $T \leftarrow T+T_l$ $k \leftarrow k+1$;
    else
      break;
5 until T_l is a leaf node;
6 repeat
  6.1 $T_r \leftarrow$The right node of T;
  6.2 If Sim(C,O) is larger than threshold τ, and $|Sim(T+T_r,O) - Sim(T,O)| < \varepsilon$
    then
      $T \leftarrow T+Tr$, $k \leftarrow k+1$;
    . else
      break;
7 until T_r is a leaf node;
8 end.

Figure 7. The outline of entities extraction algorithm.

$$\mathrm{Rel}(C, A_j) = \alpha \times \left(\frac{Itemspan(C, A_j)}{\max\{Itemspan(C, A_j)\}} \right) + (1-\alpha) \times \frac{Itemtimes(C, A_j)}{\sum_{j=1}^{n} Itemtimes(C, A_j)}$$

$$Itemspan(C, A_j) = \frac{\sum_{i=1}^{m} Span(c_i,\ A_j)}{m}$$

$$Itemtimes(C, A_j) = \frac{\sum_{i=1}^{m} Times(k_i, A_j)}{m}$$

where O is ontology *DEMLOnto*; m is properties of an application entity in *DEMLOnto*; n is properties of an entity in web page being extracted; C represents an entity semantic values set; $Span(c_i, A_j)$ is context distance between semantic C_i and ontology attribute A_j in the same page; $Itemspan(C, A_j)$ is mean value of context distance between each semantic item and the attribute A_j in the same page of an entity C. $Itemtimes(C, A_j)$ is the number of occurrences of each element and attribute A_j in an entity C. α ranges between 0 and 1, which are used to measure the influence of context distance and co-occurrence frequency on the correlation degree.

Firstly, in our extraction experiment ε is determined as 0.05 and τ is determined as 0.62 through many tests in random selected 10 web pages. Then, according to the Google ranking, we selected five top social media sites, which are mainly about books, weather and shopping based on the users' experience and the number of users. We obtain test data-set by getting certain amounts of result pages from submitting a specific query request to the selected site. Now, according to extraction process listed in Figure 7, the extraction is started from the root node of the page and the search goes down continuously to find the child node i to meet the similar condition to the inputted ontology (if the computation output of Formula 1 is large than 0.62, we judge that it is a good result). If there is no such child node i can be found we judge that there is no data can be extracted from this page record. If the node i is found, we will judge whether the sibling nodes of T_i (T_i is the sub-tree, which the node i belongs to) meets the condition or not. Finally, the sub-trees, which meet the condition, are extracted.

The evaluation index of information extraction is the accuracy rate of P, the recall rate of R and F-measure and we compare our method with traditional information extraction method, (which just uses DOM tree). The experiment result is shown in Table 3.

5. Conclusions

Currently, our ontology is developed based on manually selected classes from some baseline datasets, and our research work is limited at extraction stage. Our next target in the future is to make use of machine learning method to develop ontology from actual applications automatically and effectively. Also we will integrate ontology into opinion mining deeply.

Protégé OLV is a plug-in published in 2015. It can help us searching for relevant vocabularies in the Linked Open Data catalogue, for easing the development of ontology by reusing existing vocabularies at low fine grained level. So, taking use of Protégé OLV may largely promote the efficiency of building and reusing of the ontology. In the next, we will focus on it.

Notes

1. https://www.w3.org/TR/owl-guide/
2. http://protegewiki.stanford.edu/wiki/Main_Page

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was partially supported by the National Natural Science Foundation of China [grant number 61300170]; Key Laboratory of Computer Application Technology of Anhui Polytechnic University [grant number JSJKF201507] and The Natural Science Research Project of Higher Education in Anhui Province [grant number TSKJ2016B07].

Notes on contributors

Wan Tao received a B.S. degree in Computer Application Technology from Heihai University, China in 2004. Concurrently, she is a visiting scholar at University of Science and Technology of China. Her main research interests include web mining, semantic web, and sentiment analysis.

Tao Liu received a B.S. degree in Computer Application Technology from HeFei University of Technology, China in 2004. Her main research interests include machine learning, computer network and information security.

References

Baldoni, M., Baroglio, C., Rena, P., & Patti, V.. (2012). From tags to emotions: Ontology-driven sentiment analysis in the social semantic web. *Intelligenza Artificiale*, *6*, 41–54.

Dinakar, S., Andhale, P., & Rege, M. (2015). *Sentiment analysis of social network content. Information Reuse and Integration (IRI)*. IEEE International Conference on IEEE.

Francisco, V., Gervás, P., & Peinado, F. (2007). Ontological reasoning to configure emotional voice synthesis. Proceedings of the 1st International Conference on Web Reasoning and Rule Systems, Springer-Verlag.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition, 5*, 199–220.

Hu, C., Xu, Z., Liu, Y., Mei, L., Chen, L., & Luo, X. (2014). Semantic link network-based model for organizing multimedia big data. *IEEE Transactions on Emerging Topics in Computing, 2*, 376–387.

Liu, B., & Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In *Mining Text Data*, 415–463.

Luo, X., Xu, Z., Yu, J., & Chen, X. (2011). Building association link network for semantic link on web resources. *IEEE Transactions Automation Science and Engineering, 8*, 482–494.

Meersman, R.A. (1999). The use of lexicons and other computer-linguistic tools in semantics, design and cooperation of database systems. Star Lab Technical Report.

Noy, N.F., & Mcguinness, D.L. (2001). Ontology development 101: A guide to creating your first ontology. Stanford Medical Informatics Technical Report SMI-2001-0880. Stanford, CA.

Parrott, & Gerrod., W. (2001). *Emotions in social psychology*. Hove: Psychology Press.

Song, K., Feng, S., & Gao, W. et al. (2015). Personalized sentiment classification based on latent individuality of microblog users. Proceedings of the 24th International Conference on Artificial Intelligence, AAAI Press.

Thakor, P., & Sasi, S. (2015). Ontology-based sentiment analysis process for social media content. *Procedia Computer Science*, *53*, 199–207.

W3C Recommendation. (2009). OWL 2 Web Ontology Language quick reference guide. Retrieved from https://www.w3.org/TR/2009/REC-owl2-quick-reference-20091027/

Wu, F., Song, Y., & Huang, Y. (2015). Microblog sentiment classification with contextual knowledge regularization. Twenty-Ninth AAAI Conference on Artificial Intelligence.

Xu, L., & Lin, H., Pan, Y., Ren, H., & Chen, J.. (2008). Constructing the affective lexicon ontology. *Journal of the China Society for Scientific and Technical Information*, *27*, 180–185.

Xu, Z., Liu, Y., & Mei, L., Hu, C., & Chen, L. (2015). Semantic based representing and organizing surveillance big data using video structural description technology. *Journal of Systems & Software*, *102*, 217–225.

Xu, Z., Zhang, H., Hu, C., Mei, L., Xuan, J., & Choo, K. R. (2016a). Building knowledge base of urban emergency events based on crowdsourcing of social media. *Concurrency and Computation: Practice and Experience, 28*, 4038–4052.

Xu, Z., Wei, X., Liu, Y., Mei, L., Hu, C., & Choo, K. K. R. (2016b). Building the search pattern of web users using conceptual semantic space model. *IJWGS, 12*, 328–347.

Xu, Z., Zhang, H., & Sugumaran, V. (2016c). Participatory sensing-based semantic and spatial analysis of urban emergency events using mobile social media. *EURASIP J. Wireless Comm. and Networking*, *1*, 1–9.

Yaakub, M.R., Li, Y., & Zhang, J. (2013). Integration of sentiment analysis into customer relational model: The importance of feature ontology and synonym. *Procedia Technology*, *11*, 495–501.