基于句法依存关系的微博情感分析方法

文坤梅[1,†]　徐帅[1]

1. 华中科技大学计算机科学与技术学院，武汉 430074；

† 通讯作者，E-mail: kmwen@hust.edu.cn

摘要 本文提出了一种利用句法依存关系分析来对微博评论进行情感倾向性分析的方法。该方法首先对微博评论进行句法依存关系分析，然后按照设定的规则提取出所包含的<情感词, 评价对象>二元关系词性对，进而判断出微博评论的情感倾向性并提取出相应的评价对象。根据上述方法，实现了一个原型系统，并参加了NLP&CC 2012中文微博情感分析评测任务中的3个子任务：观点句识别、情感倾向性判断和情感要素抽取，取得不错的实验效果。

关键词 句法依存关系；微博；情感分析；

中图分类号

Microblog Sentiment Analysis Method Based on Syntactic Dependencies

WEN Kunmei[1,†]，XU Shuai[1]

1.School of Computer Science & Technology, Huazhong University of Science & Technology, Wuhan 430074;

†Corresponding Author, E-mail: wangwujun@pku.edu.cn

Abstract In this paper, we proposed a microblog sentiment analysis method based on the syntactic dependencies. Firstly, analysis the syntactic dependency of tweets, and then extract all of the binary dependency relations between sentiment words and targets based on the rules proposed in advance. Finally, identify the sentiment orientation of every tweet and the corresponding targets. According to the above method, we developed a prototype system, and participate in three subtasks of the NLP&CC 2012 Chinese Microblog Sentiment Analysis Evaluation Task, and the evaluation result is good.

Key words Syntactic dependencies; Microblog; Sentiment analysis;

1. 引言

随着互联网的发展，尤其是Web2.0 应用的普及，用户越来越倾向于在网上对各种产品或热点事件表达自己的观点。针对产品的评价无论是对于商家还是买家都十分有价值，而对于热点事件的评论则对政府做出正确的决策至关重要。观点挖掘或情感分析技术目前已在这些领域内得到大量的研究[1]。

微博作为一种新兴社交媒体，日益成为人们交流、表达观点的首选平台。它相比于传统的网络评论具有以下几个显著特点[2]：1）内容简短：微博的长度限制为140个字符，所以用户发布微博十分方便同时表达的观点容易理解；2）数据量大：微博数据的来源丰富，可以方便的通过手机、PC等终端发布；3）传播速度快：微博用户可以转发任意被关注者的微博，同时又能够使自己的关注者看到并可能再次转发，这种爆炸式的传播方式使得微博的传播速度非常快；4）实时性：微博可以通过如手机、PC、ipad等各种终端，随时随地发布出去，同时传播速度极快，这使得微博较传统的网络信息更实时。这些特点使得将微博作为用户评

论来源进行观点挖掘的研究是十分有意义的，文献[3]提出针对微博进行情感分析相对传统博客的效果将会更好，微博已经成为情感分析与观点挖掘的有效文本领域。目前针对微博的情感分析研究工作大部分的都是基于英文微博Twitter的，而对于中文微博的研究相对较少。

本文提出了一种基于句法依存关系来对中文微博进行情感分析的方法，该方法应用于NLP&CC2012中文微博情感评测任务，其中情感句识别和情感倾向性判断的正确率较高，而评价对象的抽取正确率相对较低。

2. 相关工作

情感分析或观点挖掘是指从用户主观评论文本中提取出用户对所评价对象的情感倾向性。最初有关情感分析的研究主要应用在网络在线的产品评论中，研究方法主要包括：1）基于监督的机器学习方法。Bo Pang等人将情感分析看成一个二元文本分类问题，即积极和消极两类。他们按照不同的方法提取特征，如Unigram、Bigram、POSTagging、Position等，然后对比分析SVM、ME、Naïve Bayes三种机器学习方法的实验效果，最终在IMDB影评数据集上SVM结合Unigram特征提取方法效果最好，正确率达到82%以上[4]。2）基于无监督的机器学习方法。Turney 等人在文献[5]中提出了一种基于点互信息值的方法来分析特定短语的情感极性，进而判断整篇评论的倾向性。该方法首先对待分析文本进行分词和词性标注并提取出带有形容词或副词的短语，然后基于AltaVista搜索引擎估计短语与种子情感词的点互信息值，最后根据所有短语的平均点互信息值判断文档的情感倾向性。文献[6]则将情感分为4类：喜、怒、哀、乐，并为每个类别选择了多个种子词，对中文评论文本进行了情感分析。3）基于自然语言处理技术和文本中的句子结构(n-grams)的方法。相比机器学习方法针对整篇文档进行情感分类，该方法则是细粒度到句子，即前者判断用户对整个产品或服务是否满意而后者则是用户对产品或服务的某些具体特征的态度。比如，T. Nasukawa等人针对有关数码相机的评论性文档[7]以及有关药物的评论性文档[8]利用NLP技术进行对相应产品的具体特征如产品样式、价格、功能等细粒度的用户观点挖掘。

目前随着微博的迅速发展，越来越多的国内外学者基于微博短文本进行情感分析研究。其中大部分研究工作都是基于英文微博Twitter的，比如：Barbosa[9]等人提出了一种两步分类的方法，即首先对微博进行主客观分类，然后对主观性微博进行情感分类。相对于传统主观性文本，微博评论的特征更加多样化，包括表情、超链接、图片等。Jiang等人则引入主题相关的特征进行分类，并利用微博间的转发关系来进行判断微博的情感倾向性[10]。最近，Pedro等人提出一种转换学习方法通过微博用户间的观点偏向性来推断所发布的微博的情感倾向性，其中用户间的偏向性则是通过微博转发关系以及用户间的社会关系进行推断的，该方法能够实现实时地进行情感分析。然而，在中文微博方面，研究相对较少。谢丽星等人[11]对比分析了基于表情符号的规则方法、基于情感词典的规则方法、基于SVM的层次结构的多策略方法在新浪微博数据集上进行情感分类的实验效果，结果表明基于SVM的层次结构多策略方法效果最好，并提出主题相关的特征有助于提高分类精度。

本文针对NLP&CC2012中文微博情感分析评测任务提出了一种利用句法依存关系分析来对微博评论进行情感倾向性分析的方法。该方法首先对微博评论进行句法依存关系分析，然后根据设定的规则提取出所包含的<情感词，评价对象>二元关系词性对，通过构建一个高质量的情感词库来判断情感词的情感倾向性，进而判断出微博评论的情感倾向性并提取出相应的评价对象以及对每个评价对象的情感倾向。

3. 方法介绍

在中文评论文本中，通常评价对象都是名词或名词短语，所以在本文的方法中，潜在的评价

对象限定为名词或名词短语。首先利用中科院的中文分词软件包ICTCLAS5.0 对微博进行分词和词性标注，对所有名词性短语进行按照文献[12]提出的基本名词短语的定义来进行预处理，标注出所有的潜在评价对象。然后利用斯坦福大学的句法依存关系分析工具对分词结果进行句法解析，得到两个词语间的依存关系，接着按照一定的规则提取出每个微博评论中所包含的<情感词，评价对象>二元关系词性对。通过构建一个高质量的情感词库计算每个二元词对中情感词的倾向性，进而判断出整个微博评论的情感倾向性并提取出相应的评价对象以及对每个评价对象的情感倾向。该方法的整体框图如下所示：

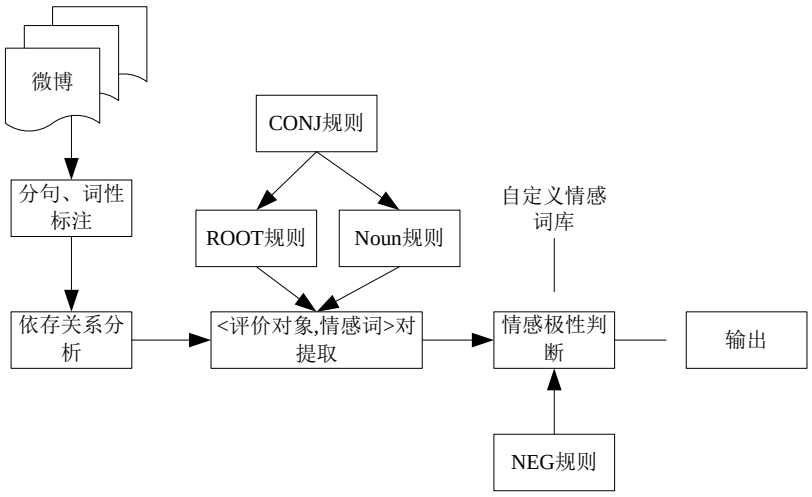

图 1 方法的整体框图

Fig.1 The overview of the method

从图 1 可以看出，该方法最核心就是<情感词，评价对象>二元关系词对的提取规则以及情感词库的构建。下面详细介绍各个步骤：

3.1 预处理

由于微博存在大量的垃圾信息，首先要对其进行一定的预处理。在本文中，主要采取了以下预处理：

1）删除微博正文中 hashtag 数目超过 3 个以上的：由于微博评论只有 140 个字，而 140 个字往往关注的话题只有 1-2 个，而 hashtag 的数目超过 3 个的往往都是广告微博或垃圾微博；

2）过滤掉微博中的链接：在微博评论往往含有一定量的超链接、图片链接、音频链接等对情感分析工作无用的信息；

3）规范化微博内容：在微博中包含一定量的英文字符以及繁体字，为了方便后续处理，将微博内容中的英文字符均转化为小写，繁体字转为简体字。同时，将所有的标点符号都转化为半角格式；另外，微博中还有大量用符号“#”表示的主题词，对于这些特殊的词汇直接作为整体标注为名词。

3.2 潜在评价对象抽取

在本文中，潜在评价对象均限定为名词或名词短语，首先利用中科院的 ICTCLAS5.0 分词工具对微博分词和词性标注。对于单个名词的词语或者满足以下条件的词组都直接提取作为潜在情感评价对象：

- BaseNP：NN|NR
- BaseNP：BaseNP+BaseNP|BaseNP+*de*+BaseNP
- BaseNP：VV+[VA+*de*]+BaseNP

– BaseNP：JJ+BaseNP

其中，BaseNP 表示基本名词短语，一般是单个名词或名词短语；*de* 表示“的|地”。

3.3 <情感词，评价对象>提取规则

本文在提取<情感词，评价对象>二元关系词对的过程中需要使用的提取规则都是基于句法依存关系来定义的。在依存文法的句法结构中，主要元素是语义依存关系，即句子中词对的二元关系，其中一个称为核心词，另一个为依存词。依存关系反映的是核心词和依存词之间语义上的依赖关系。对于分词后的微博评论，利用斯坦福大学的句法依存关系解析工具 Parser 进行句子依存关系解析，得到每个微博评论的依存关系和依存关系树。例如，句子“*卖家的服务态度不错，送货速度也很快。*”。

分词和词性标注的结果：

卖家/NN 的/DEG 服务/NN 态度/NN 不错/VA ,/PU 送货/VV 速度/NN 也/AD 很/AD 快/VA ./PU

按照 3.2 中对潜在评价对象的定义，在这里需要重构分词结果：

卖家的服务态度/NN 不错/VA ,/PU 送货速度/NN 也/AD 很/AD 快/VA ./PU

从经过重构的分词结果来看，潜在的评价对象为：卖家的服务态度和送货速度，接下来的工作便是进行句法解析，结果如下：

nsubj(不错-2, 卖家的服务态度-1) *root(ROOT-0, 不错-2)*

mmod(快-7, 送货速度-4) *advmod(快-7, 也-5)*

advmod(快-7, 很-6) *dep(不错-2, 快-7)*

其中，数字表示词语在句子中的位置，subj表示主谓关系，mod表示修饰关系，root表示与根节点的关系，详细的中文语法依存关系可参见[13]。根据中文词语语义的特点，极性词语不仅由形容词还由副词、名词、动词等词性的词语构成。因此，在本文中在提取与潜在评价对象存在依存关系的词性对时，主要考虑的词语包含以下词性标注：VA（形容词作动词）、AD（副词）、NN（名词）、VV（动词）。比如，在这个例子中，潜在评价对象“*卖家的服务态度*”和极性词“*不错*”以及潜在评价对象“*送货速度*”和“*快*”都存在直接依赖关系，如果给定一个这样的规则“当潜在评价对象和极性词间存在直接依赖关系nsubj或mmod，就认为潜在评价对象即为实际评价对象且情感倾向性由该极性词决定。”，那么就可以很容易的判断出“*卖家的服务速度*”以及“*送货速度*”都是实际的评价对象且该句为观点句，而情感倾向性则有“*不错*”和“*快*”两个极性词决定。

通过以上分析，本文参考文献[14]提出以下基于句法依存关系的<情感词，评价对象>二元关系词性对提取规则。如表 1 所示。

在表 1 中，输出结果是一个二元组的形式，第一个元素表示情感词，第二个元素表示评价对象。POS(O|T)以及 O|T_dep 分别表示词语 O 或 T 的词性和所属的依存关系。Dep(O)表示与词语 O 存在依存关系的词，{O}表示情感词集合，{T}表示评价对象集合。

Root 规则主要针对微博评论中存在大量无主谓关系的情况而设定的。在微博中，往往通过特有符号“#”来标注所讨论的主题，即评价对象，此时经过句法依存关系解析后，可以根据 root 依存关系来判断其中的依赖词是否为极性词，如果是极性词且与其存在依存关系的词语中不存在其他评价对象则将该句子的前一句的评价对象或该微博的 hashtag 作为评价对象。

利用 Target 规则提取与潜在评价对象存在特定依存关系的极性词。其中第一条规则表示极性词 O 与潜在评价对象 T 存在直接的依存关系。另外极性词可能没有直接依赖于其评价对象而是间接地通过其他词语与评价对象存在依存关系，此时可以通过第二条规则进行提取。

表 1 情感词和评价对象的依存关系提取规则

Table 1 Rules for extracting the dependency relation between opinion word and target

规则名	规则定义	规则条件	输出结果	说明
Root 规则	$root(ROOT-0, O-i)$	$POS(O) \in \{NN, JJ, VA, AD, VV\}$ $O \in \{O\}$ $Dep(O) \cap \{T\} = \varnothing$	<O, #hash#>	根据 root 依存关系中的依赖词是否为情感词来提取评价对象
Target 规则	$O \rightarrow O\_dep \rightarrow T$	$POS(T) \in \{NN, NR\}, T \notin \{O\}$ $O\_dep \in \{subj, obj, \mathrm{mod}, ccomp\}$ $POS(O) \in \{NN, VA, JJ, AD, VV\}, O \in \{O\}$	<O, T>	根据潜在评价对象所属的依存关系直接或间接的寻找与之相关的情感词
	$O \rightarrow O\_dep \rightarrow H \leftarrow T\_dep \leftarrow T$	$POS(T) \in \{NN, NR\}, T \notin \{O\}$ $O \mid T\_dep \in \{subj, obj, \mathrm{mod}, ccomp, conj\}$ $POS(O) \in \{NN, VA, JJ, AD, VV\}, O \in \{O\}$	<O, T>	

另外，为了提高评价对象提取的召回率，本文还提出基于连接关系（conj），利用已提取的评价对象来扩展待提取的评价对象的规则，如表 2 所示。

表 2 评价对象的扩展提取规则

Table 2 Extended rules for target extraction

规则名	规则定义	规则条件	输出结果	说明
Extend 规则	$T_1 \rightarrow T\_dep \rightarrow T_2$	$POS(T_2) \in \{NN, NR\}$ $T_1 \in \{T\}, T_2 \notin \{O\}$ $T\_dep = conj$	$t=T_2$	基于连接关系来扩展评价对象

在表 2 中，输出结果 t 表示通过扩展规则提取出的评价对象，并且评价对象 T2 与评价对象 T1 的评价词即情感词是一致的。

3.4 情感词倾向性判断

按照上面的规则，提取出<情感词，评价对象>二元关系词性对后，后续工作就是如何判定情感词的极性，在本文采取基于词典的判断方法。首先，以 HowNet 情感词典和 NTUSD 情感词典为基础，从中筛选出情感倾向性明确的正面情感词和负面情感词。另外，还创建了一个自定义的扩展词库主要用来扩展网络新词汇，以及考虑到微博环境下，存在丰富的表情符号以及口语化表达方式，在该扩展词库中还存放着表情符号以及一些口语化语言。

词库构建完成后，将提取出的二元词性对中的极性词，输入到构建的情感词典中，如果存在则输出其极性，如果不存在则文献[15]相似度计算软件包xsimilarity计算该极性词的情感极性，同时，为了提高情感词极性的计算效率，会将计算出来的新的情感词扩展到自定义词库中，以备下次查询。值得注意的是，在这里还需要考虑，否定关系即如果情感词被否定词修饰则需要将情感倾向性取反，依存关系分析结果中会存在依赖关系neg来标示情感词否定词间的这种关系，所以只要根据neg依存关系就可以进行否定处理即可。

4 实验结果

基于本文提出的方法实现了一个原型系统，并参加了 NLP&CC 2012 中文微博情感分析

评测任务中的 3 个子任务：观点句识别、情感倾向性判断和情感要素抽取。

4.1 任务 1：观点句识别

在本评测任务中，观点句只限定于对特定事物或对象的评价，所以如果一个句子只要能按照上述提取规则提取出一个或一个以上的<情感词，评价对象>二元关系词性对就可以判断为观点句。评测结果如表 3 所示。

4.2 任务 2：情感倾向性判断

本任务要求判断在任务 1 中判断为观点句的情感倾向。观点句的情感倾向可以分为正面（POS），负面（NEG）和其他（OTHER）。针对每个观点句，计算从中提取出的所有<情感词，评价对象>二元关系词性对中的情感词的倾向性后，如果其中既有正面词也有负面词则判断该观点句为 OTHER，否则为情感极性为 POS 或 NEG。评测结果如表 3 所示。

4.3 任务 3：情感要素抽取

本任务要求找出微博中每条观点句作者的评价对象，即情感对象。同时判断针对情感对象的观点极性。在任务 1 和任务 2 的基础上，针对每个观点句，计算与每个评价对象相关联的二元关系词性对中的情感词的倾向性便可知作者对情感对象的观点极性，评测结果如表 3 所示。

表 3 评测结果

Table 3 The result of evaluation

任务编号	微平均			宏平均		
	正确率	召回率	F 值	正确率	召回率	F 值
任务 1	0.737	0.536	**0.621**	0.743	0.522	**0.607**
任务 2	0.643	0.344	**0.499**	0.641	0.335	**0.437**
任务 3 严格评价	0.120	0.089	**0.102**	0.133	0.090	**0.105**
任务 3 宽松评价	0.198	0.138	**0.163**	0.219	0.139	**0.167**

4.4 结果分析

从评测结果来看，任务 1 的正确率较高，而任务 2 的正确率相对有所下降，说明该方法能够有效的识别出观点句，但是对于观点句的倾向性判断方面，由于基于词典的判断方法是有限的，且受到领域知识的影响很大，因此会存在很多的误判。

从任务 3 的评测结果来看，实验效果不是很理想，一方面是情感词的极性判断方面存在上述的问题；另一方面，就是在情感对象的位置定位存在很大的误差，因为微博中含有大量的垃圾信息，对其进行预处理时会对情感对象的位置造成一定的破坏，从而影响最终的实验效果。

通过以上分析，本文的方法可以考虑改进的方向为：进一步完善情感词典的构造，尤其是领域情感词典的构建。

5 结论

本文提出了一种利用句法依存关系分析来对微博评论进行情感倾向性分析的方法。该方法首先对微博评论进行句法依存关系分析，然后提取出所包含的<情感词，评价对象>二元关系词性对，进而判断出微博评论的情感倾向性并提取出相应的评价对象。通过参加 NLP&CC 2012 中文微博情感分析评测任务中的 3 个子任务证明该方法在观点句识别以及情感倾向性判别上有一定的效果，但在评价对象抽取方面效果不是很理想，需要改进的地方还有很多，尤其是情感词的完善方面。

6 参考文献

[1] B. Pang and L. Lee. Opinion Mining and Sentiment Analysis. Journal Foundations and Trends in Information Retrieval, 1-2(2):1–135, 2008.

[2] H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In Proc. of WWW2010, pages 591–600, 2010.

[3] A. Bermingham, A. Smeaton. Classifying Sentiment in Microblogs: Is Brevity an Advantage? In the Proc. of CIKM2010.

[4] Bo Pang ,Lillian Lee Vaithyanathan S．Thumbs up? Sentiment Classification using Machine Learning Techniques[C] In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Philadelphia, July 2002, pp. 79 - 86.

[5] Peter D. Turney. Thumps up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. Proceedings of 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, Pennsylvania, 2002:2

[6]段秀婷，何婷婷，宋乐．基于PMI-IR算法的Blog情感分类研究．第5届全国青年计算语言学研讨会论文集. 2010.

[7] T. Nasukawa, J. Yi. Sentiment analysis: capturing favorability using natural language processing. In K-CAP '03: Proceedings of the 2nd international conference on Knowledge capture, New York, NY, USA, 2003. ACM.

[8] J. Yi, T. Nasukawa, R. Bunescu, W. Niblack. Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In Proceeding of ICDM-03, the 3rd IEEE International Conference on Data Mining, Melbourne, US, 2003. IEEE Computer Society.

[9] L. Barbosa and J. Feng. Robust sentiment detection on twitter from biased and noisy data. In COLING, 2010.

[10] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao. Target-dependent twitter sentiment classification. In ACL, 2011.

[11]谢丽星，周明，孙茂松．基于层次结构的多策略中文微博情感分析和特征抽取．中文信息学报, 2012

[12]乔春庚，孙丽华，吴韶，王洪俊．基于模式的中文倾向性分析研究．第一届中文倾向性分析评测, 2008

[13] P. Chang, H. Tseng, D. Jurafsky, and C.D. Manning. 2009. Discriminative reordering with chinese grammatical relations features. In Proceedings of the Third Workshop on Syntax and Structure in Statistical Translation.

[14] Qiu, G., Liu, B., Bu, J. and Chen, C. 2011. Opinion Word Expansion and Target Extraction through Double Propagation. Computational Linguistics.

[15]夏天.中文信息相似度计算理论与方法[M].郑州：河南科学技术出版社.2009.