

文章编号: 1003-0077(2018)03-0049-06

基于连词的维吾尔语情感词库扩展研究

刘若兰¹, 年梅¹, 玛尔哈巴·艾赛提²

(1. 新疆师范大学 计算机科学技术学院, 新疆 乌鲁木齐 830054;
2. 新疆师范大学 文学院, 新疆 乌鲁木齐 830054)

摘 要: 极性情感词是准确分析维吾尔文倾向性的基础资源。该文在前期构建的维吾尔语褒贬情感词典基础上进行网络情感词的自动扩展研究。首先分析维吾尔语情感表达的语言特征, 总结了连词、程度副词与情感词的搭配规律, 并基于此规律设计从情感语料库中获取候选情感词的算法, 形成候选情感词库; 最后再利用维吾尔语连词的特性, 结合已创建的情感词典和维吾尔语反义词词典, 以互联网作为超大规模语料库, 设计基于搜索引擎的情感词极性判别算法, 根据算法得分判别候选情感词的极性, 再将其扩展到已构建的褒贬情感词库。实验结果表明, 与扩展前的情感词库相比, 使用互联网文本语料扩展后的情感词库后进行维吾尔语句子倾向性测评的准确率和召回率均有明显提高。

关键词: 维吾尔语; 情感词扩展; 连词; 程度副词; 极性判断
中图分类号: TP18 **文献标识码:** A

The Uyghur Emotional Lexicon Extension Based on Conjunctions

LIU Ruolan¹, NIAN Mei¹, Maierhaba Aisaiti²

(1. Department of Computer Science and Technology, Xinjiang Normal University,
Urumqi, Xinjiang 830054, China;
2. Department of Literature, Xinjiang Normal University, Urumqi, Xinjiang 830054, China)

Abstract: Emotion words are the fundamental resource for accurately analysis the opinions of the Uighur language. We investigates the automatic expansion of the web emotional words on the basis of an existing Uighur sentiment lexicon. First, we summarize the collocation rules of the conjunctions, degree adverbs and sentiment words by analyzing the linguistic features of Uighur emotional expression. Based on the rules, we design an algorithm to obtain the candidate emotional words from emotional corpus, forming the candidate sentiment lexicon. Finally, we use the Internet as a super-large corpus to design the emotional discriminant algorithm based on search engine by reusing the characteristics of Uighur conjunctions and combining with the established emotional lexicon and Uighur antonyms dictionary. The polarity of candidate emotional words is decided according to the score calculated by the algorithm, and then add them to the emotional lexicon. Compared with the emotional lexicon that was not expanded, the experimental results showed that the accuracy and recall rate of Uyghur sentence's tendency are significantly improved by our extended dictionary.

Key words: Uyghur; expansion of the emotional words; conjunctions; degree adverbs; polarity judgment

0 引言

移动互联网和社交软件的快速发展, 使网络

成为维吾尔族网民获取信息、共享资源、相互交流、发表观点的重要平台。随着基于维吾尔语的网站数量越来越多, 带有情感色彩的维吾尔语网络文本逐渐增加, 并以极快的速度广泛传播, 如

收稿日期: 2017-03-13 定稿日期: 2017-04-19

基金项目: 国家自然科学基金(61163064); 教育部人文社会科学工程科技人才培养专项(15JDGC022); 新疆师范大学数据安全重点实验室资助; 新疆师范大学计算机应用技术重点学科资助

果不加控制,网络中消极无益信息的散布将给人们带来严重危害,错误舆论的扩散甚至会干扰社会、经济的正常发展。对网络文本倾向性的快速准确分析,能够实时有效地监控舆情动态,便于政府部门及时倡导正能量,同时有利于及时过滤网络有害信息,净化网络环境,避免人们受到此类信息的危害。网络文本倾向性分析的基础资源是情感词典,情感词典的完整性和准确性将直接影响文本情感分类的准确率。由于项目组在前期构建维吾尔语基础情感词典^[1]时选用的候选情感词主要借助对中英文情感词典进行翻译和筛选形成,数量有限,尚未充分考虑网络中大量的新词和变形词。情感词的覆盖面不足,导致文本倾向性判别的精确度还不能完全满足信息过滤的需求。因此,本文在前期研究的基础上,开展基于网络语料的维吾尔语情感词库的自动扩展研究,为维吾尔语情感词的自动更新及扩展提供技术支持。

1 相关工作

通过对中英文情感词典构建的文献调研表明,现阶段情感词的获取技术主要有以下两种。

(1) 基于语义词典或知识库判别情感词,构建极性情感词典。这种方法的主要思路是:选择极性确定且明显的一组褒义词和贬义词作为种子词集,利用 HowNet、WordNet 以及同义词词林等资源提供的语义关系,设计算法获取候选词与种子词的语义相似度,以此为据判断候选词的情感倾向。例如, Kamps 等^[2]提出利用 WordNet 中词汇间语义距离的概念,计算情感倾向待选词与表示褒贬态度的基准词(“good”和“bad”)的关联度来识别词汇的情感类别。朱嫣岚等^[3]则利用 HowNet 中有关语义相似度和语义相关场的定义,设计算法得到词汇的语义倾向值,再根据语义倾向值判别词汇的褒贬倾向。黄硕等^[4]提出基于知网和同义词词林信息融合的方法,进行词汇的语义倾向计算。金宇等^[5]借助《现代汉语大词典》中词语的释义推出其他词的情感极性。

(2) 基于语料库的情感极性判别方法。该方法通过学习大规模语料中词语的搭配特征、共现和统计等特征得到一部情感词典。例如,文献^[6]利用语句中“and”、“but”等连词衔接极性形容词的特性,从语料中识别通过连词相联系的形容词对,使用图的聚类算法将抽取的形容词聚集成褒贬两类,从而实

现词汇的极性判断。王科等^[7]利用语料中的连接关系,同时结合转折词和否定词对文本倾向的影响,将语料中的词汇划分成两个词集,再进一步确定各词集对应的情感类别。Turney 等^[8]借助统计学中逐点互信息的概念衡量目标词与种子情感词的相关性,利用搜索引擎返回的 hits 数,计算目标词与所有褒、贬种子词的 SO—PMI 值来判定目标词的情感指向。阳爱明等^[9]借助 Turney 的思想构建中文情感词典,对三部开源情感词典合并去重形成基础情感词集,利用百度搜索引擎以及改进的 PMI 算法重新计算基础情感词集中全部词汇的情感权值。维吾尔语方面,文献^[10]在维吾尔语情感词汇语言特征的分析之上,设计 CRFS 的特征模板用于自动识别维吾尔语情感词汇。

目前,维吾尔语文本的倾向性分析研究尚处于起步阶段,缺乏完备齐全的维吾尔语语义资源。由于缺少可用的电子资源,基于语义词典或知识库的方法在现阶段仍难以实现。因此,本文采用基于网络文本语料库的方法,以项目组先前构建的维吾尔语情感词典资源为基础,获取包含情感词的维吾尔语情感句,分析归纳情感句中连词和程度副词与情感词的搭配特点,基于搭配关系利用网络语料抽取维吾尔语候选情感词,形成候选情感词库;最后将网络作为超大规模语料库,利用维吾尔语中并列连词联系极性相同或极性相反词汇的规律,设计了利用维吾尔语反义词词典以及 Google 搜索返回的页面结果数计算情感未知词与褒贬情感词集的相似度算法,依据计算结果判定情感类别后并入相应的褒贬词典,实现维吾尔语情感词库的不断扩展,为基于词典开展维吾尔语网络文本的倾向性分析研究提供重要工具。

2 维吾尔语情感词汇的特征

维吾尔语和其他语言相似,情感词主要聚集在名词、形容词以及它们的搭配关系中。因此,本文首先对维吾尔语情感句进行分析,总结维吾尔语中情感词的特点,然后基于连词衔接情感词、以及程度副词修饰情感词的规律从语料中抽取带有情感极性的词汇,再设计极性判别算法确定其情感类别。

2.1 连词的衔接特征

维吾尔语中,连词起衔接词与词、词组与词组、

分句与分句、句群的作用。并且连词连接的大多数句子成分是相关的,即它们表达的情感色彩可能是相同的,也可能是相反的。其中转折连词连接的句子成分通常具有相反的情感倾向,而并列连词和递进连词在大多数情况下衔接情感极性相同的句子成分,但并列连词和递进连词有时也会连接一组表达相反意义和相反情感的反义词组。

维吾尔语中,并列连词主要出现在单句中连接词汇和短语,例如, *ئۇ ھەم گۈزەل ھەم ئەقىللىق* (她既美丽又聪明),句中连词*ھەم*(既……又……)的前后分别是褒义情感词*گۈزەل*(美丽)和*ئەقىللىق*(聪明)。并列连词连接的词汇中有相当一部分都具有情感倾向,并且所连接词汇或者短语的情感极性大多数情况下相同。基于上述特点,连词既可作为候选情感词抽取的有效特征,也可借助已确定极性的情感词来判断候选情感词的语义倾向性。此外,维吾尔语中的递进连词“ئەمما”(但是)、“لېكىن”(但是)、“بىراق”(但是)、“ۋاھالەنكى”(而、然而)”等,一般用于连接句子。例如,在句子 *ئۇ چىرايلىق بۇلۇپلا قالماستىن ، يەنە مەجەزىمۇ ناھايىتى ياخشى بۇلۇپلا قالماستىن...يەنە* (她不但长得漂亮而且脾气也很好)中,递进连词*يەنە...يەنە* (不但……而且)连接了两个分句。由于表示转折和递进关系的连词连接的成分较为复杂,在现阶段用于抽取情感词过于困难,为此本文暂时未予考虑,仅选择了维吾尔语中两个使用最广泛、出现频率最高的并列连词“ۋە”和“ھەم”进行情感词的提取和极性的判断。

2.2 程度副词的修饰特征

维吾尔语情感句的表达同汉语一样,常常在情感词汇前面加上程度副词,以增强或减弱词汇原有的情感倾向强度。例如, *بۈگۈن ھاۋا بەك ياخشى* (今天的天气非常好),情感词*ياخشى*(好)被位于其前面的程度副词*بەك*(很,多么,非常)修饰,使词汇的褒义情感更强烈。针对该特点,本文总结了维吾尔语中经常与情感词一同使用的程度副词,如表1所示。将这些程度副词作为识别候选情感词的特征之一,用于提取候选情感词。

维吾尔语情感句中,还存在其他的上下文搭配关系,但由于受到现阶段维吾尔语资源和工具软件的限制,本文仅选择了上述两个特点鲜明、最易于实现的特征进行候选情感词的抽取,即连词、程度副词与情感词的搭配规则。

表 1 维吾尔语中修饰情感词的常用程度副词

程度副词	ھەددىدىن ئىزىدا (格外)	ئالاھىدە (分外)	ئىنتايىن (最)
	تامامەن (完全)	نەقەدەر (多么)	تازا (顶)
	زور (极)	شەكسىز (无疑)	قەتئىي (坚决)
	مۇتلەق (绝对)	ئالىي (极)	ئىنتايىن (十分)
	تامامەن (完全)	نېمىدېگەن (何等)	ھەقىقەتەن (的确)
	ھەددىدىن ئارتۇق (过)	ئۈزۈل-كېسىل (彻底)	ھەرگىز (绝不)
	كارامەت (棒)	پەۋقۇلئادە (极)	بەك (很)
	بارغانسېرى (越)	بولىپمۇ (尤其是)	يەنىمۇ (更)
	بەكمۇ (太)	تېخىمۇ (更)	ئۇلا (太)
	يەنىمۇ ئىلگىرلىگەنھالدا (进一步)	تولىمۇ (更)	بىر نازمۇ (一点儿也不)
	خوپمۇ (更)	زېمىن (不过)	ئارتۇقچە (过于)
	ناھايىتى (最)	ئىنتايىن (非常)	قارىغاندا (比)
	بىر قەدەر (比较)	خېلى (挺)	ھەممىدىن (比……最)
	نەسبەتەن (相对)	ھەر ھالدا (还可以)	يەنە تېخى (还)
	ئازراق (有点儿)	ئىنتايىن (不怎么)	ئىنتايىن (不那么)
	ئازغىنا (些许)	قەلچە (毫不)	بىر ناز (略微)

3 维吾尔语网络情感词的自动扩展

本文在已构建的褒贬情感词库的基础上,使用搜集的网络文本作为候选情感词的抽取语料,再以互联网丰富的数据资源作为词汇倾向性的判别语料,将语言特征和统计学方法相结合设计算法,实现维吾尔语情感词的自动扩展。

3.1 基础情感词集及其表示

本文利用项目组前期构建的维吾尔语褒贬情感词典,进行候选情感词的筛选和极性判别。目前该情感词集共包括褒义词 2 042 个,贬义词 2 473 个,对于其中的褒义词集本文用 Pwords 表示,贬义词集用 Nwords 表示,Bwords 则代表基础情感词集。本研究在此词集的基础上,基于网络语料进行维吾尔语情感词的抽取和极性判别,扩展得到的新情感词则追加到该情感词典中,以不断扩展情感词库的词汇数量和覆盖面,为文本情感分析性能的提高提供支持资源。

3.2 候选情感词识别算法

为识别维吾尔语候选情感词,本文首先根据维吾尔语情感词汇与并列连词和程度副词的上下文关系,利用收集的网络语料抽取维吾尔语候选情感词,然后借助基础情感词集从维吾尔语候选情感词集中删除极性已知的部分候选情感词。

3.2.1 语料获取及预处理

实现维吾尔语候选情感词的获取需要基于大规模语料数据。为此,本文首先使用网络爬虫工具 Hertrix 从天山网等维吾尔语网站搜集了 3 000 篇语料,其中包括新闻语料 2 580 篇和论坛语料 420 篇。其次,使用网页解析工具 HTMLParser 对收集的语料文本进行去噪处理,去除 HTML 标签,剔除无关信息,保留所需的文本内容,形成候选情感词提取的初始语料集。最后根据维吾尔语句子的标识符,进行句子划分,最终获得由 M 个句子构成的候选词提取语料集 $S, S = \{s_1, s_2, s_3, \dots, s_M\}$, 本研究中 M 为 41 176。

3.2.2 候选情感词构建算法

根据从维吾尔语情感句中总结的情感词与连词、程度副词的语言搭配规律,设计了候选情感词的抽取模板,如表 2 所示。利用此模板,结合并列连词“و”和“ھەم”以及表 1 中的程度副词,从候选词提取语料集 S 中抽取候选情感词,滤除其中出现在基础情感词集中的部分候选情感词,最终形成的候选情感词集包括 488 个词汇。

表 2 候选情感词抽取模板

序号	抽取模板	说明
1	Pword+conj+word	褒义情感词+连词+候选情感词
2	word+conj+Pword	候选情感词+连词+褒义情感词
3	Nword+conj+word	贬义情感词+连词+候选情感词
4	word+conj+Nword	候选情感词+连词+贬义情感词
5	adv+word	程度副词+候选情感词

3.3 候选情感词的极性判断算法

句子中并列连词前后的词通常具有一致的情感极性^[11],如果候选情感词经常与褒义词由并列连词衔接在一起,则该候选情感词是褒义词的可能性较大;反之,若候选情感词常跟贬义词并列相连,则候

选情感词的贬义倾向强烈。但并列连词也有连接极性相反成分的例外情况,但仅限于连接一组反义词的情形,如果候选情感词的反义词大多数为褒义词,则候选情感词的贬义情感强烈,相反,若候选情感词的反义词大部分为贬义词,则候选情感词极可能是褒义情感词。

基于以上分析,本文利用已确定情感极性的基础情感词以及维吾尔语反义词词典,分以下两种情况计算候选情感词的倾向性。在计算中,以 Bwords 表示基础情感词集, Owords 表示维吾尔语反义词词典,待判别的候选情感词为 word,其在维吾尔语反义词词典中对应 n 个反义词 Oword。

第一种情况:当候选情感词 word 的反义词出现在基础情感词集中,则可使用基础情感词集 Bwords 统计其中正负情感词汇的数量,再对统计结果取反即可获得候选情感词的极性值。

设候选情感词 word 的反义词 Oword 有 m ($m \neq 0$) 个出现在基础情感词集中,则候选情感词的极性值 $S_{\text{Value}(\text{word})}$ 的具体定义如式(1)所示,即累加这 m 个反义词的极值再取反。式(1)中 $F(\text{oword})$ 表示候选情感词的反义词的极性值,并按照式(2)对其进行赋值。

$$S_{\text{Value}(\text{word})} = - \sum_{1 \leq i \leq m} F(\text{Oword}_i) \quad \text{Oword}_i \in \text{Bwords}$$

(1)

$$F(\text{Oword}) = \begin{cases} 1 & \text{Oword} \in \text{Pwords} \\ -1 & \text{Oword} \in \text{Nwords} \end{cases}$$

(2)

第二种情况:当候选情感词 word 的反义词没有一个出现在基础情感词集中,则可以利用搜索引擎从互联网的海量数据中获取与候选情感词相联系的词汇 Cword,然后利用基础情感词集获取词汇 cword 的情感极值再进行累加,所得结果即为候选情感词的情感极值。

设候选情感词 word 共与 p 个词汇 Cword 相联系,则候选情感词的极性值 $S_{\text{Value}(\text{word})}$ 即为这 p 个词汇的极值之和,其计算方法如式(3)所示,Cword 的极性值在式中用 $F(\text{Cword})$ 表示,并根据式(4)进行计算。

$$S_{\text{Value}(\text{word})} = \sum_{1 \leq i \leq p} F(\text{Cword}_i), \quad \text{Cword}_i \in \text{Bwords},$$

$$\text{Oword} \notin \text{Bwords}$$

(3)

$$F(\text{Cword}) = \begin{cases} 1 & \text{Cword} \in \text{Pwords} \\ -1 & \text{Cword} \in \text{Nwords} \\ 0 & \text{Cword} \notin \text{Bwords} \end{cases}$$

(4)

本文通过以下方法获取 Cword: 利用能够准确

识别维吾尔文的 Google 搜索引擎,分别以“候选情感词 连词”或者“连词 候选情感词”为关键词,按精确匹配方式进行搜索,提取前 40 条搜索结果中与候选情感词通过连词直接相连的词汇,在这里本文使用“ۋە”和“ھەم”两个连词获取与候选情感词相联系的词汇。

根据以上方法赋予候选情感词情感极值后,为确定候选情感词 word 的情感倾向类别,本文设定阈值 0,按式(5)计算候选情感词 word 的情感类别。

Value(word) =
$$\begin{cases} \text{褒义词} & S_{\text{Value}(\text{word})} > 0 \\ \text{贬义词} & S_{\text{Value}(\text{word})} < 0 \\ \text{中性词} & S_{\text{Value}(\text{word})} = 0 \end{cases} \quad (5)$$

经上述算法对候选情感词进行情感倾向判别的结果见表 3,为了便于对比,表 3 也列出了计算语言学专家对候选情感词情感类别的标注结果,以及极性判别结果的准确率。经分析发现,维吾尔语网络文本资源相对于英语、汉语而言数量少,导致仅能获取少量甚至几个与候选情感词相连接的词汇,这其中又有部分词汇不是基础情感词,致使计算候选词极值时依赖的数据不够充足,极性判别结果也因此产生了相应的偏差。此外,极性判断的错误结果也源于维吾尔语、汉语的不同语法规则,维吾尔语的句子表达采用主宾谓结构,例如, ئېلىشىمىز كېرەك. بىز مۇھىتنى قوغدىشىمىز ۋە بۇلغىنىشنىڭ ئالدىنى (我们应该爱护环境并预防污染),句中连词ۋە(并)连接词语بۇلغىنىشنىڭ(污染)和قوغدىشىمىز(爱护),依据极性判别算法污染被判定为褒义词。

表 3 专家标注、极性判断算法的候选情感词 极性判别结果对比			
	专家标注	极性判断算法	准确率/%
褒义词	185	205	80.9
贬义词	239	224	86.9
中性词	64	59	83.3

为避免极性判别不正确词汇影响情感词典的质量,对本文算法的极性判别结果进行人工校正,最终得到褒义词 185 个,贬义词 239 个,将所得褒义词和贬义词分别并入相应的褒贬情感词典,最终构建的维吾尔语网络情感词典包含词汇 4 939 个,其中褒义词 2 227 个,贬义词 2 712 个。

4 实验和测试

4.1 测试语料

为验证本文提出的情感词自动扩展算法的有效

性,本文使用项目组前期构建维吾尔语褒贬情感词典时使用的测试语料进行实验,该测试语料共 2 500 句,其中正向句 1 214 句,负向句 1 148 句,无倾向性句 138 句。

4.2 测试结果

本文基于网络文本扩展维吾尔语情感词典的目的在于提高维吾尔语文本情感分类的准确率。故使用本研究扩展的网络情感词典对测试语料的 2 500 个句子进行情感分类,通过累计组成句子的词汇倾向性来确定句子的情感倾向。句子的情感值最终由句中所有情感词的情感值加和确定,本文定义褒义词的情感值为 1,贬义词的情感值为-1,为避免否定词改变句子极性对倾向性判断结果准确率的影响,记录句中否定词的出现次数,否定词出现奇数次时,句子的情感倾向值乘以-1,否定词出现偶数次时,句子的情感倾向值保持不变。最终句子的情感值大于 0 的判定为褒义,小于 0 的判定为贬义,等于 0 的则为中性。

本文通过准确率、召回率和 F 值三个指标评价测试结果。利用本文构建的网络情感词典和未扩展的维吾尔语褒贬情感词典进行维吾尔语句子极性判断的结果如表 4 所示。

表 4 基于两种情感词典的维吾尔语句子极性
判断结果对比

评测语料	词典	准确率 /%	召回率 /%	F 值
正面情感 语料	网络情感词典	72.8	65.3	0.688
	维吾尔语褒贬情感词典	65.1	57.2	0.596
反面情感 语料	网络情感词典	64.4	71.6	0.678
	维吾尔语褒贬情感词典	56.6	61.2	0.588

从表 4 可以看出,利用本文扩展的网络情感词典进行文本倾向性判别的准确率、召回率、F 值均优于使用维吾尔语褒贬情感词典的实验结果,其主要原因是本文提出基于网络语料扩展维吾尔语情感词的算法,扩展了一部分维吾尔语网络情感词和专用情感词,从一定程度上丰富了维吾尔语情感词汇,而实验使用的测试语料又来自维吾尔语网站以及微博语料的维吾尔语翻译,因此明显提升了维吾尔语句子倾向性判别的分类性能,这充分证明了本文提出的基于连词扩展维吾尔语情感词算法的可行性和有

效性。

5 结束语

情感词典作为文本倾向性分析的基础资源和重要工具,其词库的质量和数量直接影响着文本句子倾向性判别的性能。目前,维吾尔语的文本情感分析研究尚处于起步阶段,面临无任何可用维吾尔语情感词典资源的现状,项目组经过前期研究构建了维吾尔语褒贬情感词典,为了不断完善该词典,本文进行维吾尔语情感词的自动扩展研究。在总结维吾尔语情感句中搭配关系的基础上,首先依据句子中情感词与连词、程度副词的共现规律,设计候选情感词提取算法,利用搜集的网络文本语料进行候选情感词的抽取。然后继续运用维吾尔语并列连词衔接极性词汇的特征,借助已构建的褒贬情感词集、维吾尔语反义词词典和互联网的海量语料,提出利用搜索引擎获取候选词与褒贬义词的关联紧密度,并设计相应的得分算法获得候选情感词的极性值,最后根据得分判别词汇的褒贬类别。

经上述扩展算法,最终得到褒义词 185 个,贬义词 239 个,并通过实验证明了使用扩展后的情感词典进行句子倾向性判别的准确率、召回率以及 F 值均有所提高,说明本文提出的情感词扩展算法能够有效抽取并判别情感词,实现情感词的自动扩展,不断优化情感词典质量,为进行维吾尔语不同粒度的情感倾向性分析提供高质量的支撑资源。本文在情感词扩展过程中仅利用了维吾尔语情感词的连词和程度副词特征,后续将考虑运用维吾尔语感叹词以及派生词汇等多种特征进行情感词的扩展研究,进一步完善和提高情感词典的质量。



刘若兰(1991—),硕士研究生,主要研究领域为方向文本倾向性研究。

E-mail: 1273557902@qq.com



玛尔哈巴·艾赛提(1986—),硕士研究生,主要研究领域为计算语言学等。

E-mail: 278416557@qq.com

参考文献

- [1] 年梅,范祖奎,刘若兰. 维吾尔语褒贬情感词典构建研究[J]. 计算机工程与应用,2017(04): 152-155,162.
- [2] Kamps J, Marx M, Mokken R J. Words with attitude [C]//Proceedings of the 1st International Conference on Global Wordnet1. 2002: 332-341.
- [3] 朱嫣岚,闵锦,周雅倩,等. 基于 HowNet 的词汇语义倾向计算[J]. 中文信息学报,2006(01): 14-20.
- [4] 黄硕,周延泉. 基于知网和同义词词林的词汇语义倾向计算[J]. 软件,2013(02): 73-74,94.
- [5] 金字,朱洪波,王亚强,等. 基于直推式学习的中文情感词极性判别[J]. 计算机工程与应用,2011, (34): 164-167.
- [6] Hatzivassiloglou V, Mc Keown K R. Predicting the semantic orientation of adjectives[C]//Proceedings of ACL-97, the 35th Annual Meeting of the Association for Computational Linguistics, Madrid, ES, 1997: 174-181.
- [7] 王科,夏睿. 一种基于连接关系的中文情感词典构建方法[C]. 第十四届全国计算语言学学术会议, 2015.
- [8] Turney P D, Michael L L. Measuring praise and criticism: Inference of semantic orientation from association[J]. ACM Transactions on Information System, 2003, 21(4): 315-346.
- [9] 阳爱民,林江豪,周咏梅. 中文文本情感词典构建方法[J]. 计算机科学与探索,2013(11): 1033-1039.
- [10] 禹龙,田生伟,冯冠军. 维吾尔语情感词汇自动识别[J]. 计算机工程,2011(07): 213-215.
- [11] Hatzivassiloglou V, McKeown K R. Predicting the semantic orientation of adjectives[C]//Proceedings of the 35th annual meeting of the association for computational linguistics and eighth conference of the European chapter of the association for computational linguistics. Association for Computational Linguistics, 1997: 174-181.



年梅(1970—),博士研究生,教授,主要研究领域为计算机网络内容安全、信息检索研究等。

E-mail: 345566810@qq.com