

# 基于情感向量空间模型的歌词情感分析\*

夏云庆<sup>1</sup> 杨莹<sup>2</sup> 张鹏洲<sup>2</sup>

<sup>1</sup>清华大学清华信息科学与技术国家实验室, 北京

<sup>2</sup>中国传媒大学计算机学院, 北京

E-mail: yqxia@tsinghua.edu.cn; nnyy1985@cuc.edu.cn; zhangpengzhou@cuc.edu.cn

**摘要:** 音频信号在歌曲情感分析中难以奏效, 所以本文提出以歌词作为歌曲情感分析的依据, 采取基于情感单元的情感向量空间模型 (s-VSM) 进行歌词情感分析。该模型较好地解决了基于词汇的向量空间模型 (w-VSM) 在文本表示效率、歧义、情感功能和数据稀疏性等方面的不足。同时, 本文将情感词词频与 Thayer 二维情感压力模型相结合, 提出了“轻松”、“压抑”之外的“复杂”、“含蓄”两类新的情感压力类别。实验证明: (1) s-VSM 模型在歌词情感分类中优于传统方法; (2) 四类情感压力模型对歌词情感分析很有帮助。

**关键词:** 文本情感分析, 情感向量空间模型, 情绪压力

## Lyric-Based Song Sentiment Analysis with Sentiment Vector Space Model

Yunqing Xia<sup>1</sup>, Ying Yang<sup>2</sup> and Pengzhou Zhang<sup>2</sup>

<sup>1</sup>Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084

<sup>2</sup>School of Computer, Communication University of China, Beijing 100024

E-mail: yqxia@tsinghua.edu.cn; nnyy1985@cuc.edu.cn; zhangpengzhou@cuc.edu.cn

**Abstract:** Song sentiment analysis has not been satisfactorily addressed in audio signal processing community. In this work, lyric is used as proof for song sentiment analysis and the sentiment vector space model (s-VSM) is proposed to represent given lyric. Compared to the word-based vector space model (w-VSM), the s-VSM model successfully addresses the critical issues on text representation efficiency, ambiguity, functionality and data sparsity. Furthermore, the two-dimension Thayer sentiment stress model, i.e. *light-hearted* and *heavy-hearted* is extended to a four-dimension model to incorporate two extra sentiment stress levels, i.e. *complicated* and *implied*. Experiments show that 1) the s-VSM model outperforms the traditional methods; and 2) the four-dimension sentiment stress model is helpful to further improve performance of song sentiment analysis.

**Keywords:** Sentiment analysis, sentiment vector space model, sentiment stress.

### 1 前言

当前社会对歌曲的需求与日俱增, 听歌已经从个人电脑转向互联网在线视听。3G 通信网络的逐渐普及, 必然推动歌曲操作从互联网向手机扩展。为应对上述需求, 各种智能歌曲搜索和推荐系统逐渐涌现。歌曲情感分类是智能歌曲搜索和推荐的关键技术, 目标是赋予歌曲特定的情感标签, 以方便用户搜索或者系统推荐。近年来, 歌曲情感分类首先在音频信号处理研究中涌现, 人们试图借助 Mel 倒谱系数 (MFCC) 从音频信号中提取可能会反映情感的音频特征 (例如强度、

\*本文工作得到国家自然科学基金 (60703051)、科技部国际科技合作计划 (2009DFA12970) 和清华大学基础研究基金 (JC2007049) 资助。

频谱质心、能量、节奏、速度等),再借助机器学习算法实现情感分类。基于音频信号的研究已有近 20 年的历史,然而至今无法获得准确反映情感的音频特征,所取得的成效非常有限,无法达到满意的水平。考虑到目前音频信号在歌曲情感分析上的局限性,我们提出以歌词为分析对象,借助自然语言处理技术对歌曲进行情感分析。目前这方面的研究并不多见。

以歌词反映歌曲的情感存在一定风险。随着歌曲形式的不断推陈出新,歌词所表达的情感有时依赖于歌手对歌曲的演绎风格,二者甚至会发生偏差。某些歌曲从歌词看并无明显情感倾向,但经过歌手演绎后,能表达强烈的情感。调查显示:中文歌曲中只有不到 5% 的流行歌曲属于这种类型。所以,我们提出以歌词为歌曲情感的分析依据,以自然语言处理技术判定歌曲情感。

歌曲情感分析的依据是情感模型。我们采取流行的 Thayer 情感模型[1],即试图分别从能量和压力两个坐标轴划将歌曲情感分为两类,从而形成“满足(contentment)”、“沮丧(depression)”、“焦虑/狂乱(anxious/frantic)”和“生气勃勃(Exuberance)”四类情感。实验证明,音频信号在能量高低的判定上具有较高准确度,而在压力大小的判定上难以奏效。因此本文只针对压力大小的判定展开基于歌词的研究,试图从歌词中分析歌曲情感压力。这同文本观点极性分析中的“积极”和“消极”有较大相似之处。为表述方便,本文将歌曲情感压力分析简称为歌曲情感分析,两类情感定义为“轻松(light-hearted)”和“压抑(heavy-hearted)”。

本研究采取机器学习的分类方法实现歌曲情感分析,首先将歌词表示为向量空间模型(vector space model, VSM),然后以支持向量机(support vector machines, SVM)算法实现歌曲情感分类。基于词汇的向量空间模型(w-VSM)在歌词文本表示上存在如下问题:(1)尽管有很多算法可用于特征选择,但 w-VSM 无法消除某些与情感表达无关的词汇特征。这些特征不会对情感分析起到作用。(2)歌词中很多情感词汇在实际使用时存在歧义。歧义在 w-VSM 中未经适当消解而直接参与情感分析,必然对结果形成影响。(3)否定词和修饰词在歌词中频繁出现,他们对情感的增强、削弱和置反作用在 w-VSM 中无法体现。(4)歌词往往比较短,平均在 50~80 个词左右,这导致 w-VSM 严重的的数据稀疏问题。

针对上述问题,本文提出情感向量空间模型(s-VSM),以情感单元作为特征提取对象,以情感单元的统计量作为情感特征。实验结果显示,s-VSM 相对于 w-VSM 优势明显,充分证明了情感向量空间模型的有效性。

## 2 相关工作

音频信号处理研究领域在上世纪 90 年代开始歌曲情感分析研究,基本思路是以音频信号作为分析依据、采取机器学习方法进行情感分类[2,3],情感模型大都基于 Thayer 情感模型[1]。Lu 等提出层次分类方法,通过两步分析实现四类情感分类[3]。第一步借助强度特征判定能力水平,第二步借助音色和节奏特征判定压力水平。该工作也证明了音频信号在压力水平判定上的不足。

Chen 等于 2006 年开始进行基于歌词的歌曲情感分析研究[4],他们采取了类似[3]的层次分类方法。不同的是,在第二步压力水平判定上采取了歌词分析。他们采用基于词汇的向量空间模型,效果提高并不明显。Xia 等[5]提出了情感向量空间模型的初步设想,在特征定义中以情感单元取代词汇,以情感单元的统计量作为情感特征,歌曲情感分析取得显著提高。本文工作是[4]的扩展,将情感特征扩展到 12 维,覆盖了双重情感否定的情况。另外本文对情感类别的定义进行扩展,在原先“轻松”和“压抑”两类情感的基础上增加“复杂”和“含蓄”两类情感,以解决复杂情感和含蓄情感的识别。

## 3 情感向量空间模型(s-VSM)

### 3.1 基本原则

我们提出情感向量空间模型 (s-VSM) 遵循如下基本原则:

- 1) 只考虑情感相关词汇对情感分析的影响。
- 2) 情感词汇需在上下文中进行必要消歧后才用于情感分析。
- 3) 考虑否定词和修饰词对情感的置反、增强和削弱影响。

基于上述原则, 我们认为情感单元是情感特征定义的基本元素。下面我们提出情感单元的形式化定义, 并逐步给出情感向量空间模型的形式化描述。

### 3.2 形式化描述

情感词典 (L) 可描述为三元组:

$$L = \{C, N, M\};$$

$$C = \{c_i, i = 1, \dots, I; N = \{n_j, j = 1, \dots, J; M = \{m_t, t = 1, \dots, T\}$$

其中 C 代表情感关键词集, N 代表否定词集, M 代表修饰词集。这些词汇可从词典中自动获取, 每个情感词都被赋予积极或消极的极性。那么, 给定一篇歌词 W:

$$W = \{w_h, h = 1, \dots, H,$$

我们借助情感词典将 W 转换为情感单元集合:

$$U = \{u_v\} = \{c_{i,v}, n_{j,v}, m_{t,v}\}$$

$$c_{i,v} \in W \cap C; n_{j,v} \in W \cap N; m_{t,v} \in W \cap M'$$

其中  $c_{i,v}$ 、 $n_{j,v}$  和  $m_{t,v}$  出现在约定大小 (7 个词) 的上下文窗口中。实际应用中, 否定词和修饰词与情感关键词的依赖关系可通过依存分析工具获取。由于具备了情感单元这一上下文关系, 情感关键词的情感大部分歧义可被消除。基于情感单元, 我们定义如下情感向量空间模型:

$$V_s^i = (f_1^i(U), f_2^i(U), \dots, f_T^i(U)),$$

其中  $V_s^i$  代表第 i 篇歌词所对应的情感向量,  $f_j^i$  表示第 j 个情感特征。在这里, 我们采取情感单元的统计量作为情感特征。

我们根据情感关键词与否定词、修饰词的搭配关系建立 8 个情感特征, 见表 1。

表 1 本文定义的 12 个情感特征。其中  $f_{PSW}$  代表情感单元中积极情感关键词个数,  $f_{NSW}$  表示消极情感关键词个数,  $f_{NEG}$  表示否定词个数,  $f_{MOD}$  表示修饰词个数

$f_i$	满足条件的情感单元个数	$f_i$	满足条件的情感单元个数
$f_1$	$C_1 = \{f_{PSW} = 1, f_{NSW} = f_{NEG} = f_{MOD} = 0\}$	$f_7$	$C_7 = \{f_{PSW} = 1, f_{NSW} = 0, f_{NEG} = 1, f_{MOD} > 0\}$
$f_2$	$C_2 = \{f_{PSW} = 0, f_{NSW} = 1, f_{NEG} = f_{MOD} = 0\}$	$f_8$	$C_8 = \{f_{PSW} = 0, f_{NSW} = 1, f_{NEG} = 1, f_{MOD} > 0\}$
$f_3$	$C_3 = \{f_{PSW} = 1, f_{NSW} = 0, f_{NEG} = 1, f_{MOD} = 0\}$	$f_9$	$C_9 = \{f_{PSW} = 1, f_{NSW} = 0, f_{NEG} = 2, f_{MOD} = 0\}$
$f_4$	$C_4 = \{f_{PSW} = 0, f_{NSW} = 1, f_{NEG} = 1, f_{MOD} = 0\}$	$f_{10}$	$C_{10} = \{f_{PSW} = 0, f_{NSW} = 1, f_{NEG} = 2, f_{MOD} = 0\}$
$f_5$	$C_5 = \{f_{PSW} = 1, f_{NSW} = 0, f_{NEG} = 0, f_{MOD} > 0\}$	$f_{11}$	$C_{11} = \{f_{PSW} = 1, f_{NSW} = 0, f_{NEG} = 2, f_{MOD} > 0\}$
$f_6$	$C_6 = \{f_{PSW} = 0, f_{NSW} = 1, f_{NEG} = 0, f_{MOD} > 0\}$	$f_{12}$	$C_{12} = \{f_{PSW} = 0, f_{NSW} = 1, f_{NEG} = 2, f_{MOD} > 0\}$

根据情感单元的定义,  $f_{PSW}$ 、 $f_{NSW}$ 、 $f_{NEG}$  和  $f_{MOD}$  满足以下条件:

$$f_{PSW} = \{0, 1\}, f_{NSW} = \{0, 1\}, f_{NEG} = \{0, 1, 2\}, f_{MOD} = \{0, 1\} \text{ 并且 } f_{PSW} + f_{NSW} = 1$$

### 3.3 情感特征提取

我们首先利用情感词典结合依存分析工具提取歌词中的情感单元。具体过程如下：先利用情感词典在歌词中识别情感关键词，然后利用依存分析工具在约定上下文窗口中识别与该词发生依存关系的否定词和修饰词，最后实现情感单元的提取。

接下来我们分析情感单元的情感极性。若情感单元中不出现否定词，则我们简单采取情感关键词的极性作为情感单元极性。若出现了否定词，则根据否定词个数进行极性置反处理。例如双重否定将不改变情感极性。

最后我们根据表 1 所列 12 个情感特征的计算方法从歌词中提取情感特征。

### 3.4 情感向量空间模型的优势分析

我们从以下四个方面分析情感向量空间模型 (s-VSM) 相对于词汇向量空间模型 (w-VSM) 的优越性：

1) 效率：s-VSM 只考虑情感相关词并以情感单元进行特征表示，其表示效率高于 w-VSM。

2) 歧义：情感单元限定情感关键词的上下文，并结合否定词和修饰词的启发，情感歧义可在 s-VSM 模型中被大部分消除。

3) 功能：情感单元体现了否定词的置反功能和修饰词的情感增强和削弱功能，因此 s-VSM 的功能性高于 w-VSM。

4) 稀疏性：s-VSM 采用 12 个情感特征，其数目远远少于情感词汇个数，能较好解决数据稀疏问题。

## 4 基于 s-VSM 的歌词情感分类方法

本文将基于歌词的歌曲情感分析视为分类问题。在将歌曲表示为情感向量后，我们可利用训练集生成分类器，再利用分类器对歌曲进行情感分类。本文采取了性能较好的支持向量机 (SVM-light[6]) 分类方法。

最初我们根据情感压力将歌词情感划分为“轻松”和“压抑”，但实际上还存在两类之外的情感压力类别，比如“复杂”和“含蓄”。观察发现，有相当数量的歌曲在情感表达上直抒胸臆，频繁使用情感词，且“轻松”情感和“压抑”情感比例相当，表达了类似悲喜交加、又爱又恨等复杂情感。另外，少量歌曲在情感表达上文雅含蓄，很少使用情感词。我们认为，上述两类情感不能单纯归结为“轻松”或“压抑”。因此在实际处理中，我们将 Thayer 情感模型压力轴的“轻松”和“压抑”两类情感扩展为结合情感词词频的四类情感分类模型（如图 1 所示）。

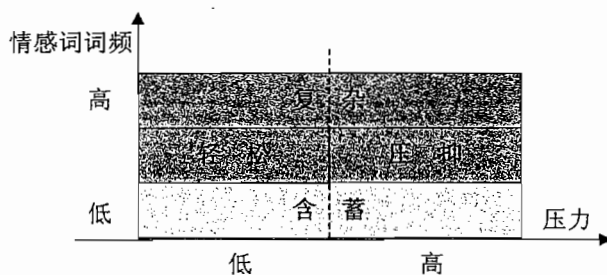


图 1 基于情感压力的四类情感分类模型

## 5 实验与评测

### 5.1 实验设置

我们采用 5SONGS 语料库[5]进行本文方法的训练和评测。5SONGS 语料库包含 2,653 首中文流行歌曲，其情感压力标签均由人工标注完成。最终 1,632 首歌曲被标注为轻松，1,021 首被标注为压抑。本实验中用到的情感词典大部分来自 HowNet[7]。由于情感词典对本文工作意义重大，因此我们又融合了 NTU 情感词典<sup>1</sup>。本文采取哈工大依存分析工具 LTP[8]进行词法分析和依存分析。

我们采取文本分类通用评测方法对本文工作进行评测，包括准确率 (p)、召回率 (r) 和 f-1 分数 (f)。为整体分析方法性能，我们采取微平均 (micro-average) 和宏平均 (macro-average) [9]。我们将 5SONGS 语料库随机划分为四等分，以四重交叉验证技术评测本文方法。

## 5.2 方法

本实验考虑如下基线系统：

### 1) 音频分析方法

采用文献[3]汇报的音频分析方法以音色和节奏等 12 维音频特征进行歌曲情感压力分析。

### 2) 知识推理方法

本文实现了一个简单的基于情感词典的情感推理方法。首先利用情感词典从歌词中识别情感词，然后在其上下文识别否定词和修饰词以获取情感单元，最后我们以情感单元计数来断定歌词的情感类别。

### 3) 基于 w-VSM 的机器学习方法

我们以情感词为分类特征，通过 CHI 算法[10]进行特征选择；以情感词特征集产生歌词向量空间；最后在训练数据上产生分类器，并用于情感分类测试。

本文方法是基于 s-VSM 的机器学习方法。我们选择 12 维情感特征产生情感向量空间，在训练数据上产生分类器，并用于情感分类测试。本实验还将对比 Thayer 的二类情感压力模型和我们的四类情感压力模型。

## 5.3 实验 1 基准方法和本文方法的对比

基于 Thayer 模型的基准方法和本文方法的实验结果如表 2 所示。

表 2 评测方法的实验结果 (f-1 分数)，其中 AB 代表音频分析方法，KB 代表知识推理方法，w-VSM-ML 代表基于 w-VSM 的机器学习方法，s-VSM 代表基于 s-VSM 的机器学习方法

方法	AB	KB	w-VSM-ML	s-VSM-ML
轻松	0.592	0.647	0.716	<b>0.768</b>
压抑	0.580	0.619	0.607	<b>0.701</b>
微平均	0.588	0.636	0.674	<b>0.743</b>
宏平均	0.583	0.625	0.662	<b>0.736</b>

表 2 实验结果显示：(1) 基于歌词的所有方法优于基于音频的方法，其中基于 s-VSM 的分类方法在微平均 f-1 分数上高于音频分析方法 0.155。这表明：在歌曲情感分析上，歌词能提供比音频更丰富的依据。(2) 基于机器学习的方法优于基于知识推理的方法，其中基于 s-VSM 的机器学习方法在微平均 f-1 分数上高于知识推理方法 0.107；(3) 基于 s-VSM 的机器学习方法优于基于 w-VSM 的方法，在微平均 f-1 分数上提高了 0.069。

## 5.4 实验 2 两个情感压力模型的对比

本文方法在 Thayer 的二类情感压力模型和本人的四类情感压力模型下实验结果如表 3 所示。需要指出的是：由于 5SONGS 语料库并未进行“复杂”和“含蓄”标注，因而无法对这两类情

<sup>1</sup> <http://nlg18.csie.ntu.edu.tw:8080/opinion/pub1.html>

感进行评测。我们只对经“复杂”和“含蓄”过滤后的歌曲进行“轻松”和“压抑”两类评测。

表 3. 本文方法在在两种情感压力模型下的实验结果 (f-1 分数)。其中 Thayer 代表 Thayer 的二类情感压力模型, Ours 代表本文提出的四类情感压力模型。

方法	Thayer (二类)	Ours (四类)
微平均	<b>0.743</b>	<b>0.831</b>
宏平均	<b>0.736</b>	<b>0.829</b>

表 3 实验结果显示: 在采用本文的四类情感压力模型后, 歌曲的“轻松”和“压抑”分类性能有显著提高 (即在微平均 f-1 分数上提高了 0.088)。这说明了新模型在歌曲情感分析上的有效性。需要特别指出的是, 本文提出的四类情感压力模型是一个面向应用的模型, 而心理学家是认同该四类情感压力模型并非本文研究重点。但我们同心理学专家合作, 以求提出反映该应用效果的新的情感压力模型。

## 6 结论

本文提出了基于情感单元的情感向量空间模型 (s-VSM)。同传统基于词汇的向量空间模型 (w-VSM) 相比, s-VSM 模型在文本表示效率、歧义消解、情感功能和数据稀疏性等方面都有 w-VSM 模型无法比拟的优越性。实验结果证明, s-VSM 模型在歌词情感分类中获得成功。另外本文对情感压力模型进行了改进, 将情感词词频与 Thayer 二维情感压力模型相结合, 提出了“轻松”、“压抑”之外的“复杂”、“含蓄”两类新的情感压力类别。实验证明, 情感压力模型的改进对提高歌词情感分析的性能很有帮助。

本研究尚有未完成的工作, 包括情感词对情感单元的增强和削弱影响尚未在本文体现, 12 维情感特征并不完整。我们将针对上述内容进一步展开我们的研究。

## 参 考 文 献

- [1] R. E. Thayer. *The Biopsychology of Mood and Arousal*, New York, Oxford University Press. 1989.
- [2] T. Li and M. Ogihara. Content-based music similarity search and emotion detection. *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing*, pp. 17–21. 2006.
- [3] L. Lu, D. Liu and H. Zhang. Automatic mood detection and tracking of music audio signals. *IEEE Transactions on Audio, Speech & Language Processing* 14(1): 5-18 (2006).
- [4] R.H. Chen, Z. L. Xu, Z. X. Zhang and F. Z. Luo. Content Based Music Emotion Analysis and Recognition. In *Proc. of 2006 International Workshop on Computer Music and Audio Technology*, pp.68-75, 2006.
- [5] Y. Xia, L. Wang, K.-F. Wong and M. Xu. Sentiment Vector Space Model for Lyric-based Song Sentiment Classification. In *Proc. of ACL-08: HLT, Short Papers (Companion Volume)*: pp.133–136, Columbus, Ohio, USA, June 2008
- [6] T. Joachims. *Learning to Classify Text Using Support Vector Machines, Methods, Theory, and Algorithms*. Kluwer, 2002.
- [7] Z. Dong and Q. Dong. *HowNet and the Computation of Meaning*. World Scientific Publishing (2006)
- [8] J. Ma, Y. Zhang, T. Liu, S. Li. A Statistical Dependency Parser of Chinese under Small Training Data. In *Proc. of IJCNLP-2004*, pp.1-5.
- [9] Y. Yang and X. Liu. A Re-Examination of Text Categorization Methods. In *Proc. of SIGIR'99*, pp. 42-49, 1999.
- [10] Y. Yang and J. O. Pedersen. A comparative study on feature selection in text categorization. In *Proc. of ICML'97*, pp.412-420, 1997.