

8-13-2020

## Friend Recommendation Based on Analysis of Users' Emotions and Text Semantics

Qun Liu

*Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;*

Hongtao Sun

*Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;*

Lianghao Ji

*Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;*

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

---

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

---

## Friend Recommendation Based on Analysis of Users' Emotions and Text Semantics

### Abstract

**Abstract:** The interests and emotions of users are often varied in their real lives. On the contrary, some other features (such as the profiles) of micro-blog are always unchangeable and they cannot describe the users' characteristics very well. Then *a novel friend recommendation method merged users' text semantics with emotions* was proposed. In the model, in order to compute the similarity of friends, some text content features from users' micro-blog are extracted and time factor was introduced. Then further consideration on the users' emotional characteristics was taken to compute the users' similarity through analyzing the emotional words in micro-blog text. Then the final results were gotten. The results of the experiments show that the model can effectively enhance the accuracy and rationality of friend recommendation by adding text semantics and emotions analysis.

### Keywords

micro-blog, text semantic, emotional analysis, similarity, friend recommendation

### Recommended Citation

Liu Qun, Sun Hongtao, Ji Lianghao. Friend Recommendation Based on Analysis of Users' Emotions and Text Semantics[J]. Journal of System Simulation, 2016, 28(11): 2852-2859.

# 一种融合文本语义和情感分析的好友推荐方法

刘群, 孙红涛, 纪良浩

(重庆邮电大学重庆市计算智能重点实验室, 重庆 400065)

**摘要:** 在现实生活中用户的兴趣和情绪波动很大, 而其他的微博特征(如注册信息等)一般变动较少, 不能很好地表征用户。提出了一种融合用户文本语义和情感分析的好友推荐方法, 根据用户的微博文本内容提取出语义特征, 通过语义分析技术来计算特征词的相似度, 同时引入了时间因素。在获得语义特征相似的用户之后, 又进一步考虑用户的情感特征, 根据微博文本中表述用户情感的词汇对用户的情感特点进行分析, 进而对上一步产生的结果做优化筛选, 得出最终的结果。通过实验表明, 加入文本语义和情感分析的好友推荐模型更能够有效地提高推荐的准确度和接受率。

**关键词:** 微博; 文本语义; 情感分析; 相似度; 好友推荐

中图分类号: TP391.3 文献标识码: A 文章编号: 1004-731X (2016) 11-2852-08

DOI: 10.16182/j.issn1004731x.joss.201611028

## Friend Recommendation Based on Analysis of Users' Emotions and Text Semantics

Liu Qun, Sun Hongtao, Ji Lianghao

(Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**Abstract:** The interests and emotions of users are often varied in their real lives. On the contrary, some other features (such as the profiles) of micro-blog are always unchangeable and they cannot describe the users' characteristics very well. Then a novel friend recommendation method merged users' text semantics with emotions was proposed. In the model, in order to compute the similarity of friends, some text content features from users' micro-blog are extracted and time factor was introduced. Then further consideration on the users' emotional characteristics was taken to compute the users' similarity through analyzing the emotional words in micro-blog text. Then the final results were gotten. The results of the experiments show that the model can effectively enhance the accuracy and rationality of friend recommendation by adding text semantics and emotions analysis.

**Keywords:** micro-blog; text semantic; emotional analysis; similarity; friend recommendation

## 引言

随着互联网技术和移动设备的迅猛发展, 用户的交友方式在发生着巨大的改变, 由以往的线下交友, 逐渐向线上交友过渡。而在线用户也希望通过网络扩大自己的交际圈, 并根据自己的需求来选择

好友, 同时通过在线社交网络平台建立起全新的沟通交流方式<sup>[1]</sup>。当前网络用户的增多使用户很难通过查找寻找志同道合的朋友来扩大自己的社交圈。以往的好友推荐算法往往会根据用户的一些注册信息、标签等不易变因素来进行用户的相似性计算, 而用户的兴趣和情绪在日常生活中, 会发生细微的变化, 需求也逐渐变得很广泛, 使得推荐算法的效果不能得到保证。

传统的推荐算法对用户进行好友推荐, 其结果往往是相互认识的朋友关系, 不能扩大其交际圈。



收稿日期: 2015-02-05 修回日期: 2015-04-08;  
基金项目: 国家自然科学基金(61379114); 重庆市自然科学基金(CSTC2014jcyjA40047); 重庆市教委研究项目(KJ1400403); 重庆邮电大学博士启动项目(A2014-20);  
作者简介: 刘群(1969-), 女, 重庆, 博士, 教授, 研究方向为智能信息处理。

<http://www.china-simulation.com>

• 2852 •

文献[2]利用标记生成树和关系图的社交网络发现好友, 考虑了用户的动态信息, 利用听音乐时的音乐标签相似性来进行好友推荐。其中采用的标签是固定的标签, 在其他方面不能得到很好的应用。文献[3]主要关注 Facebook 和 LinkedIn 上的好友推荐, 他们根据用户工作和参加相同会议的次数等行为特征来表明目标用户潜在的好友信息。这种方法只能发现同行间的好友, 却不能保证找到工作领域以外的好友。文献[4]把好友推荐转化为链路预测问题, 对用户的兴趣网络和社会网络信息进行个性化好友推荐, 这种方法的社会标签较为固定, 通常用户在选择时会非常随意不能准确表达用户的特征, 同时也没有考虑到用户在面对不同事情下情感的不确定性。文献[5]利用用户在某个地点停留时间的长短, 以及分析用户在各种网站的使用频率来判断用户的相似性以期找到共同爱好的好友。这种推荐能帮助用户找到好友, 但是位置信息只反映了用户所在的位置, 不能完全代表用户行为和喜好。文献[6]利用遗传算法对复杂网络中的拓扑特征、拓扑信息和指标进行优化来解决好友推荐的问题, 这种方法没有考虑到用户本身的内在认知信息。在现实生活中, 用户容易受到外界的影响, 各种需求较为广泛, 人类的情感认知比较丰富, 而常见的好友推荐模型都不能及时地满足用户的需求。在线社交网络中发表、转发或评论的信息, 其中不仅包含了用户的兴趣所在, 而且隐含用户的情感特征, 这些都可以作为好友推荐的一个非常重要的参考数据。

本文对用户的微博文本内容和情感进行融合分析, 建立用户的情感词词典, 通过对用户最近的微博文本进行分析, 提取出关键词和情感词。首先采用同义词转换分析技术对文本语义关键词进行处理, 得出用户的文本标签, 根据标签来计算用户的相似性。同时考虑到用户的兴趣和时间有着直接关系, 所以进一步结合了时间因素, 利用层次分析法按照时间的递进顺序对用户每天提取的文本和情感特征词分配权重。通过情感词计算用户相似度, 对第一次得到的用户相似性结果进行筛选得出

推荐结果。实验结果表明融合了用户文本和情感因素的好友推荐方法不仅能够提高推荐的准确率和合理性, 而且能够有助于用户找到具有共同兴趣和爱好的朋友。

## 1 微博文本和情感分析

### 1.1 用户行为分析

在线社交网络中的用户一般对注册信息的一些资料很少有改动, 所以单纯地利用用户的注册信息只能反映用户的过时情况, 这样的推荐就会造成偏差。随着时间的变化, 用户的兴趣标签和爱好都会发生一定的改变, 因此关注用户最近的行为最能反映用户的兴趣。以 Facebook, Tweeter 和微博为代表的在线社交网络, 用户可以发表, 评论, 转发一些短文本信息, 这些微博文本信息能实时反映用户兴趣和情感特征。

文献[7]证明在在线社交网络中不同的用户通过微博进行交互时, 能相互影响情绪。通常微博信息中会包含两类信息, 即文本信息和用户情感信息, 我们将其标记为  $mb=\{\text{Text}, \text{Emotion}\}$ , 本文给出的推荐算法能够结合用户的兴趣和情感进行分析, 实验结果表明比一般传统推荐算法更能有效产生用户感兴趣的好友。

### 1.2 用户微博文本分析

中文文本中往往包含大量的隐含信息, 而微博文本内容虽然比较短小, 但信息量却较大, 基本上是比较精简和重要的词汇。在文本相似性计算方法中首先需要从文本中提取关键词, 把提取的关键词作为用户的个性化标签, 然后计算用户的相似性以便推荐。但是在中文词语里面往往含有很多同义词、反义词以及词语相近等容易混淆的词汇, 所以针对短微博的特点, 本文基于现有的《同义词词林》对用户的文本进行语义分析。首先进行文本相似度计算, 然后再进行情感相似度计算, 所提取的用户的文本内容包含了各种各样不同的主题信息, 所以在对用户的文本内容进行相似度计算之后, 我们通

过设定阈值,从文本内容相似的用户中过滤出文本主题也相同的用户,这样可以提高推荐的计算效率,同时提高推荐的准确性。

### 1.3 用户情感分析

在《心理学大辞典》中,情感被描述为一种是否满足自己的需要而产生的态度体验。所以,在好友推荐时,是否满足用户的需求在根本上就是用户的情感分析问题。

在社交媒体中,用户的性格特征决定了他在不同的事情和问题中具有不同的情感反应,从而对其行为产生一定的影响,所以对用户情感进行深入分析才能为其找到志趣相投的朋友。微博信息不仅表达了用户对不同问题的看法,同时也包含了他们的情感信息<sup>[8]</sup>,这些都有助于对用户情感进行分析。具有相似情感的人具有一定的黏合力,对用户的情感分析已经被用在产品推荐里面<sup>[9]</sup>,具有显著的效果。

## 2 文本语义和情感分析推荐模型

### 2.1 文本相似性

在计算用户的文本相似度时,首先需要对用户  $U_i$  和  $U_j$  的微博进行向量化描述:

$$mb_i = \{(mb_{i1} = (T_{i1}, E_{i1}), \dots, (mb_{in} = (T_{in}, E_{in}))\}$$

$$mb_j = \{(mb_{j1} = (T_{j1}, E_{j1}), \dots, (mb_{jn} = (T_{jn}, E_{jn}))\}$$

用户  $U_i$  和  $U_j$  的文本相似度的交集记作  $T_{com}$ ,  $T_{com} = T_{U_i} \cap T_{U_j} = \{t_{com1}, t_{com2}, \dots, t_{comn}\}$ , 由于用户的文本内容存在时间先后的特性,距离当前时间越近,对用户的表征能力越强,为减少时间变化带来的偏差,本文在对用户的微博文本内容计算相似性时添加了相应的权重值以反映时间变化的影响。当  $T_{com}$  为空时,表示两个用户的文本不存在相似性,即文本相似度为 0,计算公式如下

$$\text{Sim}(mb_i, mb_j) = \begin{cases} \sum_{k=1}^m (wf_k \times \text{Sim}(T_{ik}, T_{jk})), T_{com} \neq \emptyset \\ 0, T_{com} = \emptyset \end{cases} \quad (1)$$

其中  $T_{ik}$  和  $T_{jk}$  分别表示用户  $U_i$  和  $U_j$  的第  $k$  天的文本内容,  $wf_k$  是表示对用户的第  $k$  条文本信息分配的权重,  $m$  代表选取文本的数量。  $\sum_{k=1}^m wf_k = 1$ , 且随着  $k$  的增大,  $wf_k$  的值越大,表示距离当前时间越近,对用户的表征能力越强。在文本相似性计算中采用 Jaccard 距离<sup>[10]</sup>, 如下

$$\text{Sim}(T_{ik}, T_{jk}) = \frac{N(T_{ik}) \cap N(T_{jk})}{\sqrt{|N(T_{ik})| |N(T_{jk})|}} \quad (2)$$

其中,  $N(T_{ik})$  和  $N(T_{jk})$  表示用户  $U_i$  和  $U_j$  的第  $k$  天微博中的关键词的集合。为了保证推荐效率,对这一步得出的待推荐列表进行过滤处理,我们选取阈值  $\delta$  为 0.4。

### 2.2 情感相似性

情感相似度的计算主要任务是提取文本中用户所产生的情感词汇。其中情感词的提取主要是基于语料库和词典的两个方法<sup>[11]</sup>。本文由于针对的对象都是中文词汇,所以采取的是基于词典的方法。课题组自定义了一个词典,对常见的程度情感词进行统计和标记,其中包含常见的程度副词 19 个,并按照程度的轻重顺序设定下标。我们通过下标的距离来计算两个情感词的相似性。若微博中含有多个情感词,在提取时我们选择下标最大的关键词,这样能比较好地反映用户情绪。与文本相似度计算类似,情感相似度计算也考虑了时间的因素,其中情感词词典构成的向量为:  $\text{Emotion}_{\text{dict}} = \{em_1, em_2, \dots, em_i\}$ , 则两个用户的情感相似性度计算方法如下

$$\text{Sim}(U, U_i) = \sum_{k=1}^m (wf_k \times \frac{1}{1 + \alpha |E - E_i|}) \quad (3)$$

其中  $\alpha$  是距离衰减参数,当两个情感程度词的距离越远,表明两者的相似性越小,反之亦然。若两个用户的情感词在同一位置,则相似度为 1,若距离较远,则需要参考  $\alpha$  的取值进行计算。在得出最后的好友推荐列表时,需要对列表进行过滤,以保证推荐的质量,同文本相似度的阈值取值类似,选取阈值  $\varepsilon$  为 0.4。

## 2.3 文本语义和情感分析推荐模型

本文所提出的文本语义和情感分析推荐(SER)模型框架如图 1 所示。

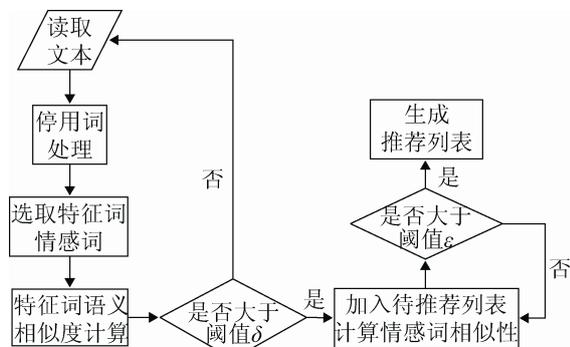


图 1 SER 模型

通过公式(2)可以计算出两个用户的文本相似性, 得出第一个待推荐列表。由于不同的用户在表述事情存在不同情感的差异, 所以需要继续对用户的情感进行分析, 在微博文本中经常包含了一些情感词汇, 我们根据自定义的情感词典, 提取出用户的情感词, 然后计算情感相似性, 从而过滤第一个推荐列表。

融合文本语义和情感分析的好友推荐算法步骤如下所示:

算法 1. 基于用户文本语义和情感的好友推荐算法

输入: 用户  $U$  的 ID

输出: 推荐列表 RM

步骤:

1. 初始化  $R_m \text{List} = \emptyset$ ,  $R_m = \emptyset$ ;
2. 提取用户  $U$  的文本特征词和情感特征词  $\{(T_1, E_1), (T_2, E_2), \dots, (T_5, E_5)\}$ ;
3. 用户集合  $U_{\text{list}} = \{U_1, U_2, \dots, U_n\}$ ;
4. 分别提取用户集合中用户的文本和情感特征词:  $\{(T_{i1}, E_{i1}), (T_{i2}, E_{i2}), \dots, (T_{i5}, E_{i5})\}$ ;
5. FOR  $U_i \in U_{\text{list}}$  //遍历用户列表
6.  $\text{Sim}(U, U_i) = \sum_{k=1}^m (wf_k \times \text{Sim}(T_k, T_{ik}))$
7. IF  $\text{Sim}(U, U_i) > \delta$

8. 把  $U_i$  添加到  $R_m \text{List}$  中
9. END IF
10. END FOR
11. FOR  $U_i \in R_m \text{List}$  //遍历待推荐列表
12.  $\text{Sim}(U, U_i) = \sum_{k=1}^m (wf_k \times \frac{1}{1 + \alpha |E - E_i|})$
13. IF  $\text{Sim}(U, U_i) > \varepsilon$
14. 把  $U_i$  添加到 RM 中
15. END IF
16. END FOR
17. 返回 RM

## 3 实验验证与结果分析

### 3.1 实验环境及数据

实验运行环境是 jdk1.8, Eclipse 和 mysql5.5。实验数据的获取是通过自行编写的爬虫程序进行抓取, 在爬取数据时我们采取随机爬取策略, 随机选择若干个微博用户, 爬取这些微博用户的好友信息, 根据好友的 ID 再依次爬取这些好友的好友微博等信息, 然后重复此过程。分别爬取了两组新浪微博数据, A 组数据是 2014-05-03-2014-05-17 的数据, B 组数据是 2013-11-13-2013-11-22 的数据, 数据中包含了用户的注册信息, 微博数据信息, 好友关系和微博关系。如表 1 所示。

表 1 爬取数据集的统计信息

数据特征	A 组数量	B 组数量
用户数	63 641	6 083
微博数	84 168	10 569
用户关系	1 391 718	80 692

在微博网络中, 由于存在很多僵尸粉和组织机构账户, 这样的用户虽然拥有很多的粉丝, 但是数据集中若含有大量的该类用户会对实验结果产生一定的影响, 所以对数据进行长尾分布验证。图 2, 3 展示了采集的数据集中用户节点的出度和入度分布情况, 表明了节点度数呈现长尾分布特征, 也说明所采集数据集符合复杂网络中节点度的幂律分布特点<sup>[12]</sup>, 数据集是有效和真实的。

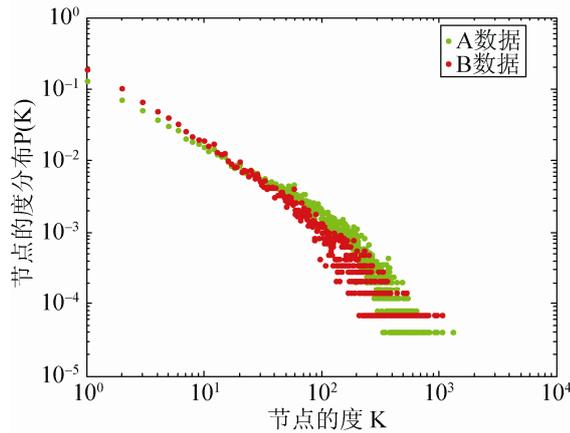


图2 用户节点出度的概率分布图

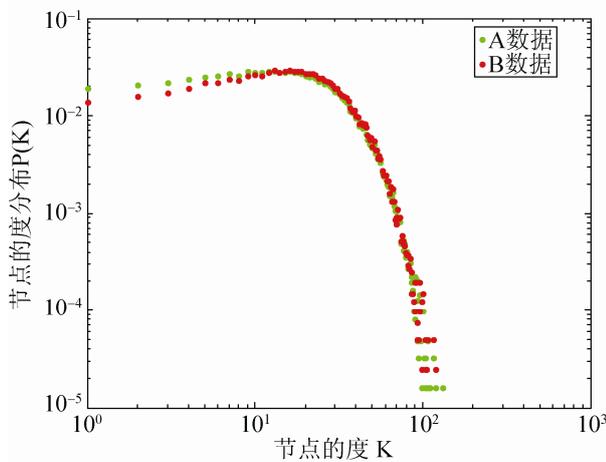


图3 节点入度的概率分布图

### 3.2 评价指标

推荐系统中,最常用的指标是采用 Top-K 推荐中的准确率、召回率,同时为了考虑推荐结果中用户顺序,加入了评价指标  $P@N$ 。

准确率可以表示为能够正确推荐出的好友个数与推荐好友总数的比值,计算方法如公式(4)。

$$\text{Precision} = \frac{N_{hit}}{N_r} \quad (4)$$

式中:  $N_{hit}$  表明推荐好友是用户真实好友的个数;  $N_r$  是总共推荐的好友数量。

召回率即查全率,可表示为能够正确推荐出的好友个数与用户好友总数的比值<sup>[13]</sup>,计算方法如公式(5)。

$$\text{Recall} = \frac{N_{hit}}{N_A} \quad (5)$$

其中:  $N_{hit}$  同上;  $N_A$  是用户的好友总数量。

$P@N$  本身是  $\text{Precision}@N$  的简称,指的是对特定的查询,考虑位置因素,检测前  $N$  条结果的准确率。计算公式如下

$$P@N = \frac{T_N \cap R_N}{N} \quad (6)$$

其中:  $T$  是根据抓取的前后顺序生成的用户列表;  $T_N$  为前  $N$  个好友;  $R_N$  是给用户推荐的前  $N$  个好友。

### 3.3 计算文本和情感权重

在经济学中的层次分析法可以用来解决多因素下的决策问题,所以本文引用层次分析法 (Analytic Hierarchy Process, AHP) 对每个因素(本方法中分别是微博文本分析中产生的不同词汇和微博文本产生的时间)确定权重值<sup>[14]</sup>,该方法可以解决复杂多目标下的定性与定量相结合的问题。主要分为两大步骤:构建判断矩阵和一致性检验计算。

#### 3.3.1 构建判断矩阵

在判断矩阵的构建过程中,指标的确定是以五分位标度为基础<sup>[15]</sup>,在用户的微博文本和情感相似性公式中,权重  $wf_k$  的计算由 AHP 进行计算得出。为了保证推荐的实时性,每次计算都是以用户的 5 条微博为基准,构建的判断矩阵如表 2 所示。

表2 微博文本的判断矩阵

微博文本	第1条	第2条	第3条	第4条	第5条
第1条	1	3	5	7	9
第2条	1/3	1	3	5	7
第3条	1/5	1/3	1	3	5
第4条	1/7	1/5	1/3	1	3
第5条	1/9	1/7	1/5	1/3	1

#### 3.3.2 一致性检验计算

构建判断矩阵是考虑其中的客观成分是否足够合理化,由于客观成分相对复杂并且人们的认识比较主观,所以需要对判断矩阵进行一致性检验。检验指标包括一致性指标(CI)和一致性比率(CR),其公式如(7)所示。

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (7)$$

$$CR = \frac{CI}{RI}$$

其中:  $\lambda_{\max}$  是判断矩阵的最大特征根;  $n$  是比  $\lambda_{\max}$  小的最大整数;  $RI$  是随机一致性指标, 当一致性比率  $CR < 0.1$  时, 认为构建的判断矩阵, 可以作为权重参与计算。通过公式(7)可以计算得出  $\lambda_{\max} = 5.2375$ ,  $n = 5$ ,  $CI = \frac{\lambda_{\max} - n}{n - 1} = 0.0593$ ,  $RI = 1.12$ , 则  $CR = \frac{CI}{RI} = 0.053 < 0.1$ , 符合一致性条件。计算的权重结果如表 3。

表 3 实验权重设置

$wf_1$	$wf_2$	$wf_3$	$wf_4$	$wf_5$
0.543	0.226	0.109	0.065	0.057

### 3.4 实验结果

为了更好地验证算法的有效性, 分别对 A、B 两组数据进行对比试验。实验结果分别考虑了在文本相似度的基础上加入情感词和不加入情感词。实验中所采用的对比方法有以下几种。

1) FOF+, 根据对共同的好友关系, 计算得出好友推荐结果<sup>[16]</sup>。

2) FOF+Tag (Friend of Friend), 此方法是根据用户的共同好友进行标签相似性匹配<sup>[5]</sup>。

3) FOF+BP (friend of friend plus birthplace), 这种方法对共同的好友和注册信息中的出生地进行推荐<sup>[17]</sup>。

4) UMR(Unified Micro-blog Friend Recommendation), 此方法中结合了标签, 位置, 签到, 和热点话题的信息进行推荐<sup>[18]</sup>。

在对用户的情感信息进行分析时, 利用公式(3)来进行计算。在实验之前需要对  $\alpha$  进行取值分析, 通过 F-measure 的变化来选取  $\alpha$  的最佳值。实验如图 4 所示。

从图 5 中可以看出当  $\alpha = 0.5$  时, 好友推荐的效果能够做到最好。所以, 在对用户的情感词相似性计算中, 取  $\alpha$  为 0.5, 代入式(3)可得到公式(8)。

$$Sim(U_i, U_j) = \sum_{k=1}^m (wf_k \times \frac{1}{1 + 0.5 \times |E_i - E_j|}) \quad (8)$$

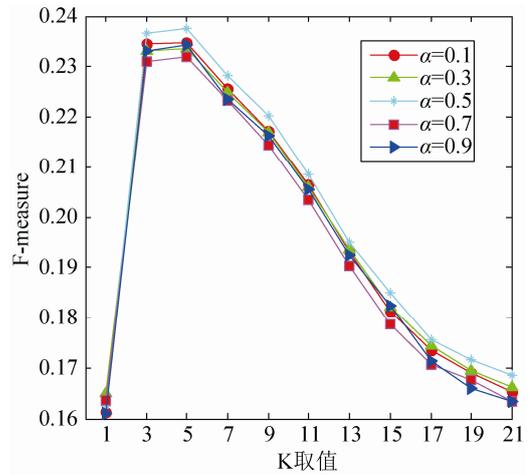
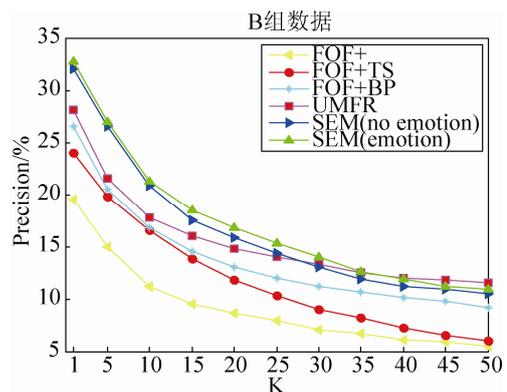
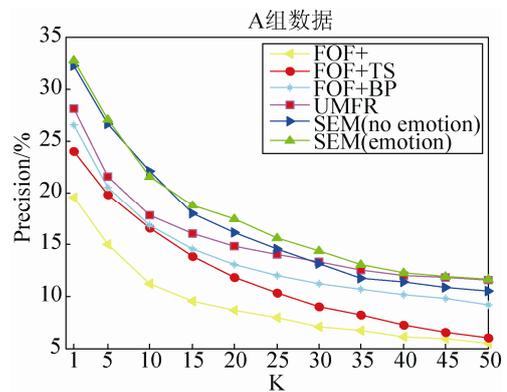
图 4 不同  $\alpha$  值下的 F-measure 值比较

图 5 精准率

首先在原数据中, 利用随机分层的采样策略进行采样。根据用户的好友个数进行层次划分, 然后对每层的用户进行随机采样, 我们对采样的用户进行实验, 对得出结果求平均得到最终的精准率、召回率和 P@N 值。两组数据的精准率, 召回率,

P@N 值的实验结果如图 5, 6, 7 所示。

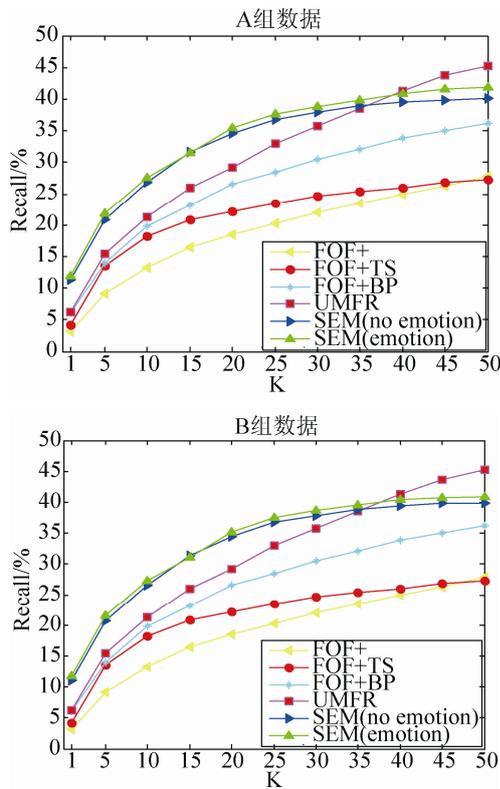


图 6 召回率

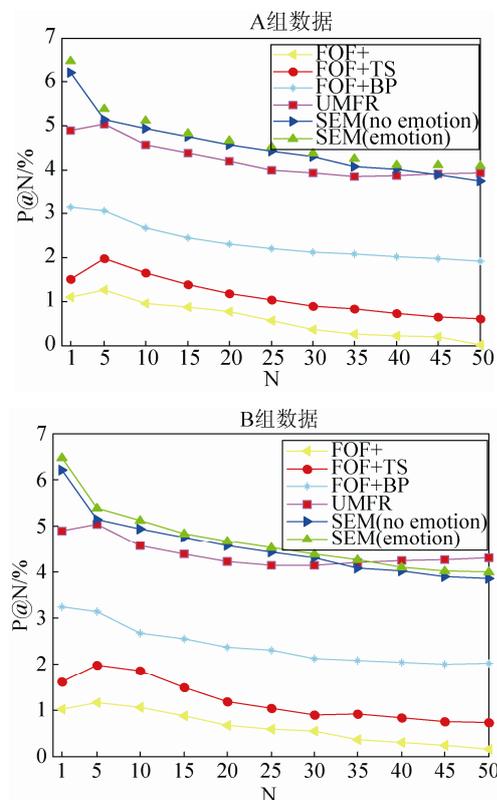


图 7 P@N 的结果

从上述 3 组图对实验结果分析如下:

1) 图 5 是表示的 A、B 两组数据下的精确率对比图,从图中可以看出,本文提出的算法在推荐的个数少于 35 时,所产生的推荐效果较高。而且加入情感分析之后,精确率有明显的提高。

2) 图 6 是表示 A、B 两组数据下的召回率对比图,从图中可以看出,本文提出的算法在推荐个数少于 40 时,能保持很好的结果,加入情感因素同样能够提高召回率。

3) 图 7 是表示 A、B 两组数据的 P@N 的计算结果,从图中可以看出,本文提出的算法在得出的 P@N 值较其他算法效果更好,而且加入情感分析之后,比不考虑情感因素得到略微提高,随着推荐好友个数的增加,趋于平稳。

通过上述实验,可以得出对用户的情感分析,在一定程度上可以提高推荐的精确率和召回率,但在不含情感词的试验中,本文提出的推荐模型会随着推荐用户的增多而有所下滑。然而在在线社交网络中,用户能够记住的好友数目符合“邓巴数字”这一规律,提供过多的好友反而造成用户的负担,这说明本文提出的推荐模型能够充分提高用户的接受度和满意度,并且方便地帮助用户找到兴趣相同的朋友,扩大自己的交际圈。

## 4 结论

本文提出了一种通过在线社交网络中的微博文本信息融合文本和情感分析的好友推荐模型,模型通过对用户微博文本进行关键词和用户情感词的提取,以语义分析技术为基础对特征词进行转换,再利用 Jaccard 算法对其特征词进行相似性计算得出待推荐列表,在此推荐列表基础之上,加入情感相似性对上面待推荐列表进行过滤,最终将筛选后的结果推荐给用户。通过实验表明,融合了用户文本语义和情感分析的好友推荐能够比传统的好友推荐算法的精确度有所提高,且能实时的对用户的兴趣和爱好进行分析,提高推荐的质量,在一定程度上可以提高推荐的新颖性。

## 参考文献:

- [1] Yin Z, Gupta M, Weninger T, et al. Linkrec: a unified framework for link recommendation with user attributes and graph structure [C]// Proceedings of the 19th international conference on World Wide Web. USA: ACM, 2010: 1211-1212.
- [2] Gou L, You F, Guo J, et al. Sfviz: interest-based friends exploration and recommendation in social networks [C]// Proceedings of the 2011 Visual Information Communication-International Symposium. USA: ACM, 2011: 15.
- [3] Chin A, Xu B, Wang H. Who should I add as a friend?: a study of friend recommendations using proximity and homophily [C]// Proceedings of the 4th International Workshop on Modeling Social Media. USA: ACM, 2013: 7.
- [4] Yang T, Cui Y, Jin Y. BPR-UserRec: a personalized user recommendation method in social tagging systems [J]. The Journal of China Universities of Posts and Telecommunications (S1005-8885), 2013, 20(1): 122-128.
- [5] Chu C H, Wu W C, Wang C C, et al. Friend recommendation for location-based mobile social networks [C]// Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2013 Seventh International Conference on. USA: IEEE, 2013: 365-370.
- [6] Silva N B, Tsang I R, Cavalcanti G D C, et al. A graph-based friend recommendation system using genetic algorithm [C]// Proceedings of the 2010 IEEE Congress on Evolutionary Computation. USA: IEEE, 2010: 1-7.
- [7] Xiong X B, Zhou G, Huang Y Z, et al. Dynamic evolution of collective emotions in social networks: a case study of Sina weibo [J]. Science China Information Sciences (S1674-733x), 2013, 56(7): 1-18.
- [8] 赵妍妍, 秦兵, 刘挺. 文本情感分析 [J]. 软件学报, 2010, 21(8): 1834-1848.
- [9] Bao S, Xu S, Zhang L, et al. Joint emotion-topic modeling for social affective text mining [C]// Data Mining, 2009. ICDM'09, Ninth IEEE International Conference on. USA: IEEE, 2009: 699-704.
- [10] Liben-Nowell D, Kleinberg J. The link - prediction problem for social networks [J]. Journal of the American Society for Information Science and Technology (S1532-2882), 2007, 58(7): 1019-1031.
- [11] da Silva N F F, Hruschka E R, Hruschka E R. Tweet sentiment analysis with classifier ensembles [J]. Decision Support Systems (S0167-9236), 2014, 66: 170-179.
- [12] ME N M G. Finding and evaluating community structure in networks [J]. Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics (S2470-0045), 2004, 69(2): 292-313.
- [13] Shi Lei, Han Yingjie, Ding Xiaoguang, et al. An SPN-based Integrated Model for Web Prefetching and Caching [J]. Journal of Computer Science and Technology (S1000-9000), 2006, 21(4): 482-489.
- [14] Ho W. Integrated analytic hierarchy process and its applications—a literature review [J]. European Journal of Operational Research (S0377-2217), 2008, 186(1): 211-228.
- [15] Saaty T L. How to make a decision: the analytic hierarchy process [J]. European Journal of Operational Research (S0377-2217), 1990, 48(1): 9-26.
- [16] Hsu W H, King A L, Paradesi M S R, et al. Collaborative and Structural Recommendation of Friends using Weblog-based Social Network Analysis [C]// AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs. USA: AAAI, 2006: 55-60.
- [17] Kamishima T, Akaho S, Asoh H, et al. Efficiency Improvement of Neutrality-Enhanced Recommendation [C]// Decisions@ RecSys. Hong Kong, China. USA: ACM, 2013: 1-8.
- [18] Feng S, Zhang L, Wang D, et al. A Unified Microblog User Similarity Model for Online Friend Recommendation [M]// Natural Language Processing and Chinese Computing. Germany: Springer Berlin Heidelberg, 2014: 286-298.