

微博话题符号网络下的意见领袖挖掘算法研究*

曹林林, 郑明春[†]

(山东师范大学 管理科学与工程学院, 济南 250014)

摘要: 在现有的意见领袖的挖掘中主要是将网络结构和用户行为作为研究方向,没有考虑到是否真正对用户产生影响这一重要的属性。针对意见领袖的挖掘进行进一步的研究,使用符号网络作为研究工具,通过赋予用户之间的观点关系链接相应的代表支持或者反对的符号,将传统的意见领袖挖掘算法结合符号网络中能够描述用户观点变化的符号关系,将真正对用户产生影响的意见领袖挖掘出来,从而挖掘得出更加精准有效的意见领袖。

关键词: 社交网络; 意见领袖挖掘; 符号网络

中图分类号: TP301.6

文献标志码: A

文章编号: 1001-3695(2017)12-3547-05

doi:10.3969/j.issn.1001-3695.2017.12.006

Algorithm of opinion leader mining based on signed network

Cao Linlin, Zheng Mingchun[†]

(College of Management Science & Engineering, Shandong Normal University, Jinan 250014, China)

Abstract: With the rapid development of mobile Internet, user gradually become the leader of social media, the abrupt rise of new media has changed the traditional information's dissemination pattern and regularity. In the existing mining of opinion leaders, it is mainly from the research of network structure and user behavior without considering an important attribute: whether the user has a real impact. Counter the mining of the opinion leaders, this paper took the symbolic network as the research tool. By giving symbol which correspondingly represented support or oppose to the link about point of view relationship between users and combining traditional algorithms of mining with symbolism which could describe the change of view between users, this paper got the opinion leader who had real impact on users. The result shows the proposed algorithm is more accurate and effective.

Key words: social network; opinion leader mining; signed network

0 引言

意见领袖最初是由拉扎斯菲尔德在一次选举中发现提出的,源于二级传播理论,表达了信息产生从媒介到大众的过程中,一小部分人对信息的解读对其他人产生的影响,并提出信息的传播是从大众媒体—意见领袖—大众的一个传播过程。在这个传播过程中,意见领袖发挥着重要作用。在新的 Web 2.0 时代,网络赋予个体不同的传播途径,使得每一个个体都有机会去主动传播个人的观点,同时拥有一呼百应的能量,那么在新的信息时代,意见领袖的定义^[1]是否出现新的解读?意见领袖发挥作用的方式、发挥的作用是否发生变化呢?本文通过对新时代的意见领袖的解读,通过分析意见领袖的特点发现新的挖掘切入点,并挖掘出更加有效、精准的意见领袖,进而对网络中的舆论传播、控制、方向的引导发挥一定的作用。

符号网络^[2]的研究最初起源于社会学领域,是指边具有正或负符号属性的网络,其中正或负符号分别表示积极或消极的关系。以符号网络去分析理解复杂网络中的拓扑结构具有十分重要的理论意义。科学技术的发展和复杂网络研究的兴起为符号网络带来了新的机遇,不管是在社交网络还是在电子商务平台,庞大的用户群中总会存在相近或对立的观点,

符号网络研究的关键在于综合利用正负边信息传达用户之间的观点,并且对个性化推荐、用户分析和态度预测都有很重要的应用价值。信息技术的快速发展提高了符号网络研究的准确性,同时在社会、生物等其他领域也出现了新的发展趋势。

本文以符号网络为工具,将意见领袖的挖掘和符号网络相结合,通过符号网络中的符号来更加准确地表示用户之间的影响关系,使得最终挖掘得出的意见领袖更加精准、更加有效。本文以新浪微博的数据为实验数据,与已有的工作相比,本文的主要贡献在于:a)通过建立一个面向微博话题的用户关系符号网络,从用户静态属性、行为特征和观点变化方面挖掘用户的影响力,综合考虑了用户在特定话题领域的领导作用;b)提出一种基于符号网络的话题意见领袖挖掘算法,针对特定话题的用户符号网络,在传统的 PageRank 算法的基础上扩展网络中的属性特征,从基础影响力、话题影响力以及观点影响力方面来挖掘意见领袖,提高了话题意见领袖挖掘结果的精准度。

1 相关工作

在传统的意见领袖定义中,意见领袖是属于广义上的受众群体,多为社会知名人士、技术专家、名人和明星,他们拥有媒体的话语权,能够对公众产生一定的影响。然而由于互联网的

收稿日期: 2016-09-09; 修回日期: 2016-10-28 基金项目: 国家自然科学基金资助项目(61402266); 国家社会科学基金资助项目(14BTQ049)

作者简介: 曹林林(1992-),女,硕士研究生,主要研究方向为社交网络意见领袖挖掘; 郑明春(1963-),女(通信作者),教授,硕士,博士,主要研究方向为社交网络、计算机网络信息安全、服务质量管理等(zhmc163@163.com)。

普及、网络的快速发展,一定程度上提升了意见领袖的影响力,而网络的虚拟环境在一定程度上削弱了现实社会地位对意见领袖影响。意见领袖最重要的价值在于作为一种信息的传播媒介,而不是信息源。因此,从这一层面而言意见领袖是指在信息的传播中,意见领袖并不一定是信息的首发者,因而信息的首因效应不是那么明显,通过意见领袖的转播,信息在社交网络中扩散的速度加快;意见领袖在传播信息的同时会加入自己对该信息的理解以及解释,从而对其他的受众产生一定程度的影响。

目前关于网络意见领袖挖掘的算法主要分为四类:a)用户属性分析法^[3-5],根据用户的静态属性,即粉丝数、微博数、转发数等来进行判别,忽略了网络的动态性以及用户的交互属性等;b)信息交互分析法^[6-8];c)网络结构分析法^[9,10],通过计算节点的节点中心性、度中心性、介数中心性等网络结构属性来挖掘意见领袖,准确度有一定的局限性;d)基于主题的挖掘^[11],将微博中的用户具有主题依赖性考虑到挖掘算法中。在意见领袖的定义中已经提到,意见领袖的一个重要属性就是能够对其他的用户产生影响,这种影响具体是什么影响不容易明确地定义,但基本上可以从几个方面来表达,就是意见领袖产生的影响是影响其他用户的观点,是一个用户对某一问题从未知到已知、从陌生到熟悉、从错误到正确的改变,这些改变可以从从用户的评论文本中得出。

数学问题研究中有一种重要的研究工具即符号网络,本文将之应用到意见领袖的挖掘中,可以直观地反映出用户的观点变化。在现有的研究中,意见领袖挖掘主要考虑到同质性^[12],而支配性也起到很大的作用。支配性主要是在对其他用户的控制力上,具体体现在对用户的观点变化的影响上,综合考虑用户的结构属性、静态属性以及用户的观点变化,可对意见领袖的挖掘更加地准确,找出的意见领袖更加地有效。

2 意见领袖挖掘框架

本文将社交网络微博的某一话题下的所有用户节点之间的边定义为有评论活动交流,两节点之间连线,边的属性具有正负性。正负属性的获取是由用户之间的评论关系得出的,通过对文本内容的切分,提取出极性词语进行极性判断,从而得出用户之间的关系属性为正或者负,这是用户成为意见领袖一个决定性的因素;同时关注用户的静态属性(包括粉丝数、发博数等)以及用户的交互属性(转发数、评论数等)。本文基于对以上用户特征的分析,利用文本内容分析、提取文本符号^[13],在社交网络中的微博某一话题下对用户的影响力进行计算,进一步得出意见领袖用户。

本文总体技术框架如图1所示,主要包括以下内容:

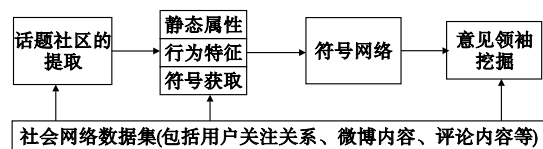


图1 意见领袖挖掘总体框架

a)提取出微博中的某一特定话题的用户节点属性,形成需要研究的话题社区网络。

b)对节点的特征进行分析,这里的节点特征包括用户的静态属性和用户的交互属性,同时对用户之间评论的文本进行

极性分析,得出边的正负属性,构建符号网络。

c)用户的影响力计算。在分析用户的节点属性的基础上,根据意见领袖挖掘算法计算出用户的影响力排名。

3 用户的特征分析

3.1 话题社区提取

在微博这个社交网络中,用户的行为具有一定的话题聚集性,同一兴趣爱好的用户会由于某一话题而参与到共同话题的讨论中,不同的用户发表各自的观点,同时还会对其他用户的观点进行适当的评论,从而形成话题社区。话题社区本身的特点为用户节点的特征提取和分析提供了丰富的支持,有助于意见领袖挖掘算法的设计和实现。

话题社区提取是进行意见领袖挖掘的前提,因为意见领袖具有话题依赖性,由于社会资本的差异性,导致不同的用户在不同的话题方面具有不同的话语权,所以本文在不同的话题中找出的意见领袖会有很大的差异。而本文中话题的选取具有一定的要求,选择具有争议的话题,用户的观点具有比较清楚的区分,使得实验的效果更加明显。

本文通过新浪微博提供的API接口进行数据的获取,获取的数据集中包括用户的静态属性和交互属性,还有一个重要的指标就是数据中带有时间节点,这为进行不同时间段的观点获取提供了可能。获取话题社区以后,挖掘用户影响力特征成为意见领袖节点识别的关键因素,用户节点的影响力是多种因素综合影响共同作用的结果,基于对用户节点的综合分析,本文选取静态属性、交互属性和符号变化作为影响用户影响力的因素。话题社区的评论关系如图2所示。



图2 话题社区的评论关系

3.2 静态属性和交互属性分析

静态属性体现了用户本身具有的特点和所在网络中与其他用户的基本关系,包括用户的粉丝好友数和体现用户活跃度的微博数、对他人微博的评论数量。交互属性包括用户在该话题社区发表的相关微博数量、被评论的数量和被提及的数量,在一定程度上反映用户的影响力。由于用户是通过文本对其他的用户产生影响,用户之间的评论微博的数量越多,用户之间的交流就会越多,互相产生影响的概率就会越大。

将用户的静态属性和交互属性量化,得出属性量值。对于静态属性而言,用户的节点权重与用户的静态属性相关,具体

的计算公式按照原有的研究如式(1)所示。其中 N_1 表示用户的相关博文数量; N_2 表示用户的粉丝数量; N_3 表示用户的评论数量,这里的评论数量是指该用户对其他用户的微博文章进行评论; $\beta_1 \sim \beta_7$ 分别表示相应的子属性所占的比重,它们的和为 1。边 e_{ij} 的权重 $w_{ij} \in W$, 与用户 v_i 与 v_j 的交互属性转发数量、被评论数量、被提及数密切相关,具体按式(2)计算。其中 N_4 表示用户发表微博的相应被转发数; N_5 表示用户的博文的相应被评论数; N_6 表示用户发表的博文受到的相应被提及数; $\sigma_1 \sim \sigma_3$ 表示对应的属性所占的比重。

$$p_i = \beta_1 N_1 + \beta_2 N_2 + \beta_3 N_3 \quad (1)$$

$$w_{ij} = \sigma_1 N_4 + \sigma_2 N_5 + \sigma_3 N_6 \quad (2)$$

由于所得到的数据具有不同的参考标准,不能直接进行比较和判断,为了方便比较和有效地评价,需要把各个指标进行归一化的处理^[14],最终把数值变换到 $[0, 1]$ 上,本文用式(3)对指标进行归一化。

$$Y = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (3)$$

3.3 符号提取

在意见领袖的挖掘中,本文不能简单地从静态属性和交互属性入手进行影响力的排名计算来衡量用户是否为意见领袖,还需要从意见领袖的定义入手,是否对其他用户产生影响是一个很重要的影响因素,因为在定义中,意见领袖的观点是会影响他人的观点。本文从用户之间的评论文本内容入手,分析用户对待观点的态度以及一个用户对待另一个用户的针对该话题的看法的态度,在社交网络中,特别是微博这个言论较为自由的社交网络,原创微博或者说热门话题会引发大量的用户参与,带来大量的评论内容,这些评论体现了不同用户的不同的认知理解,综合体现了用户的支持或者敌对态度。因此本文将微博评论内容作为态度分析的对象。

本文采用基于情感词典的机器学习方法对微博评论进行内容极性分析。内容的极性分为正向极性、负向极性和中性,为了便于研究,对获取的实验数据进行预处理,将一些噪声数据和明显的不带有极性词语、没有用户态度的数据删除,同时将一部分死节点去除掉,保留对本文实验有用的部分数据。因此本文构建的符号网络是一个不完全符号网络,称之为有限符号网络。

本文采用的微博评论内容极性的特征如表 1 所示,使用经过标记的 250 000 条正向带有极性的微博文本和 250 000 条带有负向极性的微博文本并基于决策树方法训练分类器,采取 10 次十折交叉验证的方法保证分类器的性能。根据决策树分类模型,对话题社区的参与用户之间的微博评论内容进行极性分析,得到用户之间的观点关系,构建出用户之间观点关系的符号网络。本文对所获取的数据集极性根据时间点的不同进行两次分析,第一个时间段 t_1 分析之后,再对下一个时间段 t_2 的数据进行极性分析。

获得符号网络之后,本文对来年各个时间段的符号网络进行对比,可以得出用户之间的观点随着互动的进行而发生变化。本文对这种变化进行量化,设用户 $u_j (j=1, 2, \dots)$ 对用户 u_i 发表的微博进行了评论,将用户 u_i 所获得的所有的正向极性的评论的比例定义为用户 u_i 的共鸣比例:

$$e_{ij} = \frac{N_{\text{pos}}}{N_{\text{total}}} \quad (4)$$

其中: N_{pos} 表示 u_j 对 u_i 的微博评论中正向极性的评价数目; n_{total} 表示用户 u_i 获得的总的评论数目。

表 1 用于内容极性分析的特征

特征	描述
1	正向极性词个数
2	负向极性词个数
3	否定词
4	程度副词
5	带有明显极性的标点符号个数
6	正向表情符号个数
7	负向表情符号个数

4 符号网络构建

根据微博话题数据集对数据进行整理,将用户节点为死节点或者用户之间的关系极性分析后为没有明显极性关系的用户节点不作显示。将用户之间的评论内容进行极性分析,得到用户之间的观点关系,构建出有限的符号网络。抓取的数据集是带有时间节点的数据,本文利用用户之间的交流互动,分析不同的时间段用户的观点受影响而变化的情况。在 t_1 时间得到的话题社区的符号网络如图 3 所示;在 t_2 时间得到的话题社区的符号网络如图 4 所示。

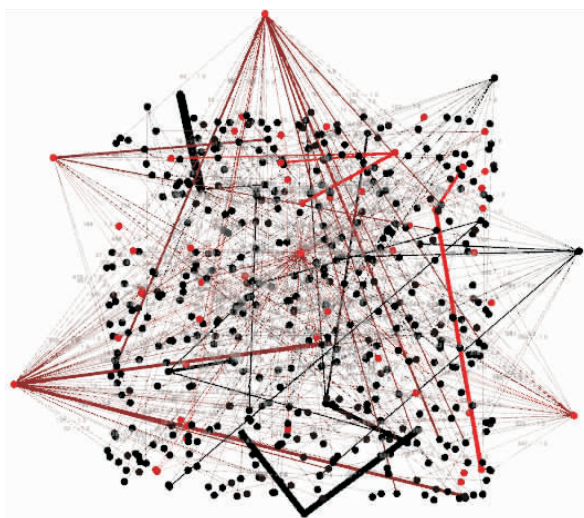


图3 t_1 时间段话题社区符号网络

在本文的符号网络中,用红色的边代表用户之间的观点关系为反对关系,黑色的边代表用户之间的关系为赞同关系,参见电子版。在构建的前后两个不同时间段的符号网络中,通过统计计算得出不同的时间段用户之间的观点变化(本文中采用计算用户的正向评论的比例的差值作为参考数据),作为用户影响力的影响因素之一进行用户的影响力计算。

5 意见领袖挖掘算法

在构建好的特定话题下的用户关系符号网络中,挖掘意见领袖就转换为寻找图中心节点的问题。传统的图中心节点的挖掘算法中,如普通的复杂网络中心性计算方法和最初的 PageRank 算法^[15],均未考虑到图中的节点的权重和边的权重问题,因此并不适用于进行话题下的意见领袖挖掘,得不到精确的意见领袖。本文通过对 PageRank 算法的改进,在微博的话题中挖掘意见领袖节点,综合考虑用户节点的静态属性、用户行为特征和符号特征,提出了意见领袖挖掘算法 InfluenceRank

算法,算法使用式(5)来计算用户节点的影响力值。

$$IR(u_i) = d \times (p_i + \sum_{j=1}^N w_{ij}) + (1-d) \times \sum_{j \in \text{comj}} \frac{\theta_{ji}}{\theta_{jh}} IR(u_j) \quad (5)$$

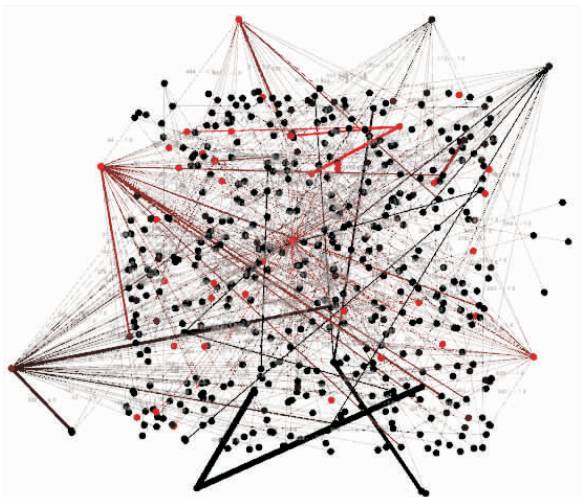


图4 t_2 时间段话题社区符号网络

其中: $IR(u_i)$ 代表节点*i*的影响力值; p_i 表示用户的节点权重; w_{ij} 表示用户的边权重; d 是阻尼系数; θ_{ji} 表示用户 u_j 对用户 u_i 的评论权重。

本算法是由 PageRank 算法改进而来的,在现阶段的研究中,对 PageRank 算法的改进已有不少的文献,其中较有代表性的吴颀辉等人^[16]提出的采用 TopicLeaderRank 算法进行意见领袖的挖掘,该算法利用用户的社交属性和内容属性,并综合用户的交互信息构建用户的行为网络,没有涉及到内容层面的分析。但是在本文的算法中,综合考虑了用户的静态属性、交互属性,同时在交互属性中加入了决定用户是否为意见领袖的一个重要的因素,即是否对他人的观点产生影响,通过对文本的内容极性分析提取出用户的观点极性,生成用户之间观点关系的符号网络,并通过符号的变化得到直观的表达。

由于用户的社会资源、社会价值观的不同,对待问题各有各的独特的看法,在社交网络中,网络的自由度比较大,用户可以自由地发表言论,对于自己不认同的观点可以直接表示反对,对于与自己的认知相符的表示支持;通过用户之间的互动,可以改变他人的认知或者对待某事件的看法,从而成为一个较有影响力的用户。该算法加入了用户之间的评论互动极性,得到能够表达用户之间观点关系的符号网络,所以在本算法中加入这一影响因素,也就是加入用户的共鸣比例,从而进行用户影响力的计算,使挖掘的结果更加精准。

在上述算法中提到的参数问题,本文采取 AHP 确定。AHP 是一种实用的多准则决策方法,以其定性和定量相结合处理各种决策因素。本文采用 Delphi 法通过比较确定每个层次内指标的相对重要性,根据重要性的等级,因素 1 相对于因素 2 的重要性一般可分为同等重要、比较重要、绝对重要,对应的标度可用 1,3,5,7,9 表示,处于两个判断中间的状态用 2,4,6,8,表示。对指标进行比较判断后,运用 AHP 法得到判断矩阵,然后进行一次性检验,可以得到各评价指标权重 w ,结果如表 2 所示。

表2 评价指标及一致性检验

指标	发博数	粉丝数	评论数	转发数	被评论数	被提及数
权重	0.4	0.2	0.4	0.35	0.4	0.25

注:一致性检验: $CR=0.005 < 0.1$

6 实验分析

6.1 实验数据与设置

本文采用来自新浪微博的热门话题数据集作为实验数据,包括 2014 年两会期间的关于雾霾和房价的讨论,详细信息如表 3 所示。为了保证微博用户和微博信息的话题相关性,数据的获取包括两个阶段:a)从科研数据共享平台数据堂(<http://www.datatang.com>)下载其他用户整理后上传的相关话题的原始微博话题数据;b)通过新浪微博开放平台的官方 API 获取这些微博的转发、评论以及涉及的用户信息。

表3 数据集信息

话题	原创微博数	转发微博数	评论数	用户数
两会雾霾	6 997	7 731	7 527	17 151
两会房价	13 111	35 880	35 105	67 464

本文的实验中,将 InfluenceRank 算法与仅考虑用户属性节点权重排序算法和仅考虑用户间交互关系的 Opinionleader-Rank 算法^[16]进行了对比实验,其中节点权重排序算法是本文设置的一个基准算法,它仅对于所有的参与话题的用户按照节点权重从大到小排序,然后将排名靠前的用户当做意见领袖。由于原始的 OpinionleaderRank 算法仅考虑用户间的评论关系,而本文构建的用户行为网络不仅考虑了用户的评论关系和转发关系,还将用户之间的共鸣性产生的影响力考虑进来,所以实验时对算法进行了改进和扩展,即在同时考虑用户之间的转发关系和评论建立边,还考虑到不同用户的影响力是不同的。

6.2 实验结果

本实验中,将排名前 10 的用户作为意见领袖,表 4 列出了 InfluenceRank 和 TopicLeaderRank 算法在“两会雾霾”话题下得出的意见领袖的相关特征。

从表 4 可以看出, InfluenceRank 算法得出的意见领袖在平均质量上都要优于 TopicLeaderRank 算法,它找出的意见领袖在整体方面的表现都比较优秀。从结果中可以看到两种算法所得到的意见领袖基本是不同的,重合度非常小,但是考虑了用户的观点变化的算法明显地要比不考虑观点变化的算法的结果更加地精确,可能有的用户的其他基本属性不是很大,比如说“北京绿林创新”,该用户的粉丝数和评论数不是很大,但是该用户与其他用户的互动比较频繁,能够对其他用户的观点产生影响,从而得出的影响力值比较大,排名相对就比较靠前。

在本文的实验研究中发现用户的共鸣比例分布特征,在统计的实验数据中,对两个话题进行分析得出用户的共鸣比例如图 5 所示。从图 5 中可以看出用户的共鸣比例在 0.4~0.6 的用户数量最多,同时用户之间的共鸣比例非常小或者非常大的用户特别少,基本符合正态分布的特征。

覆盖率与核心率从网络结构和用户交互的角度来评价意见领袖。本文比较了两种算法在话题中找出的意见领袖,图 6、8 展示了意见领袖值排名前 20 位用户的覆盖度情况,图 7、9 展示了意见领袖值排名前 1% 的用户的核心率情况。从图中可以看出,在覆盖率和核心率这两个评价指标上两个算法的差异不是很大,因为这两个指标都是从用户交互的角度来评价意见领袖,但是 InfluenceRank 算法在核心率上还是要稍微优于 TopicLeader-Rank 算法,但是覆盖度要略劣于该算法。虽然覆

盖度与核心率都未涉及用户的属性,但是同时考虑用户属性、用户交互的影响力和网络符号特征的 InfluenceRank 算法要比考虑用户属性和用户交互的 TopicLeaderRank 略优。

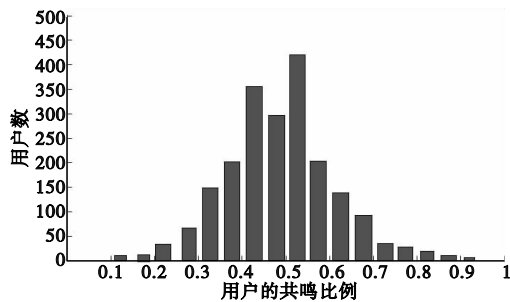


图5 话题社区中的用户共鸣比例分布

表4 两种算法在“两会雾霾”话题下排名前10的意见领袖的信息

算法	排名	昵称	微博数/评论数/粉丝数	转发数/被评论/提及数
TopicLeaderRank 算法	1	人民日报	25421/4/17649830	1085/5803/56
	2	头条新闻	85728/0/31197126	457/5124/31
	3	央视新闻	28568/1/15892195	450/4001/15
	4	南方都市报	34150/2/25825222	250/3543/12
	5	协和章荣娅	6487/3/1404917	102/3148/10
	6	张醒生	17099/0/507094	98/1407/9
	7	天涯社区	7313/2/824611	85/863/9
	8	人民网	37735/0/11791709	93/644/7
	9	陆琪	10917/0/20897119	89/512/5
	10	陈里	15993/0/20055352	36/203/3
InfluenceRank 算法	1	人民日报	25421/4/17649830	1085/5803/56
	2	头条新闻	85728/0/31197126	457/5124/31
	3	话题	3723/121/1277016	201/4522/45
	4	爱可丽防雾	819/45/11743	193/1476/27
	5	霍口罩 KN95	1317/26/124	82/876/21
	6	北京格林创新	15993/31/1203	52/423/15
	7	JEEP2700	13837/42/2459	41/201/10
	8	黄浦工商	XZ5568/23/2587	61/62/6
	9	心想事成	1633/12/645	17/23/5
	10	长沙市曙光环保公益发展中心	1963/8/25185	9/18/3

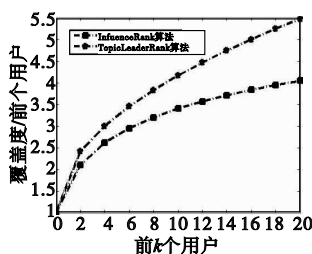


图6 两种算法对于“两会雾霾”的覆盖度

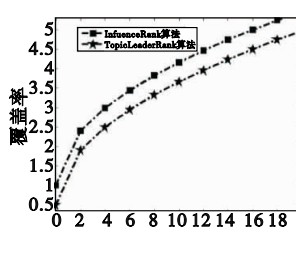


图7 两种算法对于“两会雾霾”的核心率

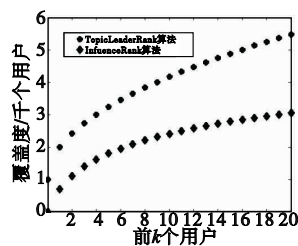


图8 两种算法对于“两会房价”的覆盖度

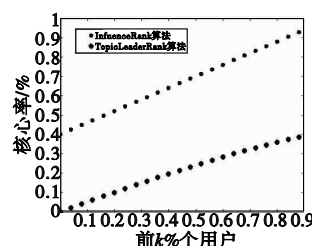


图9 两种算法对于“两会房价”的核心率

在支持率方面,本文算法得出的意见领袖的支持率也是相对较高的。支持率是所有的用户对该用户的正向的评论占所有评论的比重,根据本文的算法所得到的意见领袖是得到大部

分用户认可的权威用户或者是初始没有得到认可、但随着互动的过程使得对方的观点发生变化,最终与自己的观点趋于一致的具有影响力的用户。

7 结束语

本文对微博中的话题意见领袖挖掘进行了研究,结合用户的静态属性、交互属性和符号网络特征,提出了一种新的挖掘算法。该算法更加地注重意见领袖的影响,以用户和意见领袖之间的交互作为重要的因素进行研究,加入了用户的观点变化等重要的影响因素进行意见领袖的挖掘。本文在真实的微博数据基础上对算法进行实验验证,结果表明该算法是行之有效的,得出的意见领袖更加精准、更加具有说服力。但是本文还存在一些不足,在下一阶段的研究中,会加入更多的时间节点进行分析,从而得到更加详细具体的实验结果分析。

参考文献:

- [1] Li Feng, Du T C. Who is talking? An ontology-based opinion leader identification framework for word-of-mouth marketing in online social blogs[J]. *Decision Support Systems*, 2011, 51(1): 190-197.
- [2] Cheng Suqi, Shen Huawei, Zhang Guoqing, et al. Survey of signed network research[J]. *Journal of Software*, 2014, 25(1): 1-15.
- [3] 丁雪峰,胡勇,赵文,等. 网络舆论意见领袖特征研究[J]. *四川大学学报:工程科学版*, 2010, 42(2): 145-149.
- [4] 尹衍腾,李学明,蔡孟松. 基于用户关系与属性的微博意见领袖挖掘方法[J]. *计算机工程*, 2013, 39(4): 184-189.
- [5] 张伟哲,王佰玲,何慧. 基于异质网络的意见领袖社区发现[J]. *电子学报*, 2012, 40(10): 1927-1932.
- [6] 吴岷辉,张晖,赵旭剑,等. 基于用户行为网络的微博意见领袖挖掘算法[J]. *计算机应用研究*, 2015, 32(9): 2678-2683.
- [7] Agaerwal N, Liu Huan, Tang Lei, et al. Identifying the influential bloggers in a community [C]//Proc of International Conference on Web Search and Web Data Mining. New York: ACM Press, 2008: 207-218.
- [8] 樊兴华,赵静,方滨兴,等. 影响力扩散概率模型及其用于意见领袖发现研究[J]. *计算机学报*, 2013, 36(2): 360-367.
- [9] Cho Y, Hwang J, Lee D. Identification of effective opinion leaders in the diffusion of technological innovation: a social network approach [J]. *Technological Forecasting and Social Change*, 2012, 79(1): 97-106.
- [10] Page L, Brin S, Motwani R, et al. The paper rank citation ranking: bringing order to the Web[R]. Stanford: Stanford Infor Laboratory, 1999.
- [11] 冯时,景珊,杨卓,等. 基于 LDA 模型的中文微博话题意见领袖挖掘[J]. *东北大学学报:自然科学版*, 2013, 34(4): 490-494.
- [12] 邱泽奇,范志英,张树沁. 回到连通性——社会网络研究的历史转向[J]. *社会发展研究*, 2015(3): 1-31, 242.
- [13] HowNet. How Net's home page[EB/OL]. <http://www.keenage.com>.
- [14] 王连芬,许树柏. 层次分析法引论[M]. 北京:中国人民大学出版社, 1990.
- [15] Miao Qingliang, Zhang Shu, Meng Yao, et al. Domain-sensitive opinion leader mining from online review communities[C]//Proc of the 22nd International Conference on World Wide Web. 2013: 187-188.
- [16] 吴岷辉,张晖,杨春明,等. 一种话题相关的微博意见领袖挖掘算法[J]. *小型微型计算机系统*, 2014, 35(10): 2296-2301.