

特定话题传播网络中的意见领袖检测方法^①

兰 天^{1,2,3}, 郭躬德^{1,2}

¹(福建师范大学 数学与计算机科学学院, 福州 350007)

²(福建师范大学 网络安全与密码技术福建省重点实验室, 福州 350007)

³(龙岩学院 网络信息中心, 龙岩 364012)

摘 要: 针对中文微博目前已有的意见领袖识别模型存在的不足, 提出一种特定话题传播网络中的意见领袖检测方法. 识别模型通过转发关系建立信息传播网络, 以用户自身权威值和转发用户的支持力来评价用户的影响力. 通过对微博两周以来特定话题下意见领袖的检测实验, 结果表明该方法能够有效地识别特定话题下的意见领袖.

关键词: 意见领袖; 舆情监控; 传播网络; 识别模型

Opinion Leader Detection Method in the Communication Network of the Specific Topic

LAN Tian^{1,2,3}, GUO Gong-De^{1,2}

¹(School of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350007, China)

²(Network Security and Cryptography Key Laboratory of Fujian Province, Fujian Normal University, Fuzhou 350007, China)

³(Information and Network Center, Longyan University, Longyan 364012, China)

Abstract: For the problems of the deficiency of existing recognition model of the Chinese weibo opinion leaders, one detection method of opinion leaders in a dissemination network of the specific topic is proposed. A recognition model based on the forward relations on the basis of information dissemination network is proposed in this paper, and the influence of users comes from their own authority and the support of forwarding user. An experiment of weibo opinion leader detection for about two weeks which is concentrated in a certain topic verifies that this method can effectively detect the opinion leader of the specific topic.

Key words: opinion leader; public opinion monitoring; communication network; recognition model

1 引言

在 web 2.0 时代对互联网发展的极大推动下, “数据网络”这一概念普遍深入人心, 人们的日常生活和互联网相结合已经成为这个时代的明显特征. 在《第 34 次中国互联网络发展状况统计报告》^[1]中显示, 截止 2014 年 6 月, 中国网民用户群已达 6.32 亿, 其中手机用户群已达 5.27 亿. 国际上, twitter 平台在 2006 年兴起并不断流行和推广, 该平台已成为访问量最大的网站之一. 国内与之对应的新浪微博一枝独秀, 而微博客这一信息即时传播、用户关系分享的社交平台, 已成为中国网民上网的主要活动之一. 同时新浪微博也应用了很多策略来达到他的领袖地位, 例如邀请了政府组织、媒体机构、名人名流等地加入, 使得

重大信息的发布都变得尤其迅速. 据 2015 年微博发布的第三季度财报中显示, 截止 2015 年 9 月 30 日, 微博月活跃用户数已经达到 2.12 亿人. 同时, 微博也已经成为社会舆论传播、社会风气导向的重要地点, 因此微博舆情监控便成为国家政府和网络监管部门面前的一个大课题.

微博舆情监控的主要手段是找到舆论传播中的“意见领袖”. “意见领袖”这一概念最早由美国传播学者 Lazarsfeld^[2]在 1948 年提出, 他认为意见领袖在信息传播中扮演着对他人传递信息并施加影响的作用. 而后 Rogers^[3]在研究中表示, 意见领袖同时具有“信息中间人”和“有影响力的人”两种角色. 在消息流通如此迅速的社交网站中, 意见领袖更是扮演着举足轻重的作

① 收稿时间:2016-03-26;收到修改稿时间:2016-04-29 [doi:10.15888/j.cnki.csa.005485]

用。近几年,国内对意见领袖的研究逐渐升温,主要研究领域在于论坛、博客、学术论文、微博等。随着新浪微博成为社交领域的领头地位,意见领袖的研究已成为一大课题。

微博中用户通常针对某一主题发布微博,表达自己的见解和看法。而其他用户可以通过转发行为来进行信息的扩散,并通过转发关系构建了一个信息网络。随着信息网络中的用户逐渐提高形成了舆论群体,其中部分用户成为意见领袖,引导着舆论的发展方向和传播速度。因此本文通过特定主题下的转发关系,结合网页排序算法来识别信息网络中的意见领袖,同时验证该算法的有效性。

2 相关工作

目前应用于微博意见领袖识别的方法主要有:①基于网页排序算法进行改进,使其适用于微博信息网络中对用户影响力的排序;②基于用户特征提取影响力指标,建立评分模型;③基于消息传播网络,对网络节点进行分析评价。

基于网页排序算法,主要是对网页排序中传统的PageRank^[4]和HITS^[5]算法进行改进。Weng^[6]根据twitter中的用户结构和微博主题之间的相似度,提出了TwitterRank算法,使PageRank算法有效地移植到twitter中。Xiao^[7]根据中文微博特定结构,构建了收听网络和信息转发网络,并根据传播行为特征提出了类PageRank的WeiboRank算法。Yuang^[8]利用被关注度替代粉丝数来消除微博中虚假粉丝数的影响,重新定义影响因子,并结合PageRank对用户影响力进行评价。Lu^[9]根据特定话题下的用户个体特征和转发特征,引入时间衰减函数,提出了PageRank的改进意见领袖模型(IOLM)。Xiong^[10]根据微博信息转发关系,构建了意见领袖网,将HITS算法应用于用户权威值和中心值计算,提出了HITS-BOWR算法。

在评分模型构建方面,主要分析影响力相关属性,对特定参数进行提取。Liu^[11]从用户影响力和用户活跃度两个角度考虑,提出了使用层次分析法和粗糙集决策分析法对意见领袖特征进行识别,并提取决策规则,最后在研究中发现意见领袖是主题依赖的。Wang^[12]在研究中得出用户的关注量、粉丝量、身份认证和发布的微博量这四个方面的意见领袖识别的关键,并以这四个方面的参数建立了意见领袖识别模型。Li^[13]提出以活跃度、传播力和覆盖度三个指标来评估微博意

见领袖的影响力,构建评估指标体系,并利用层次分析法和属性特征权重排序得到影响力的最终值。Ding^[14]综合考虑了转发关系、回复关系、复制关系、阅读关系,提出了基于多关系网络的随机游走模型MultiRank,并将用户分为“多话题层次影响力个体”和“单话题层次影响力个体”。Wang^[15]根据意见领袖在信息传播中难以量化表示的问题,提出了一种基于消息传播的微博意见领袖影响力建模与测量分析方法,可以定量地对传播过程的初始影响力、影响力衰减指数及其影响力持续时间等指标进行评价。

在基于消息传播网络的研究中,主要在转发关系的基础上进行研究,分析节点的影响力。Zhao^[16]根据影响力扩散模型(IDM)存在的缺陷,引入了有效关键词概念,并对信息设置影响因子,提出了一种新的影响力扩散概率模型(IDPM),更加有效地评价影响力。Zhou^[17]针对传统意见交互模型的构建环境是封闭的社交网络,提出一种基于意见领袖引导作用的网络舆论演化分析方法。Zhang^[18]根据区分微博中用户的转发行为分为“主题相关转发”和“跟随转发”两种关系,指出被转发概率高而具有高影响力的用户不一定是专家,提出了一种基于主题模型的概率生成模型。

3 背景知识

PageRank^[4]是一种根据网页之间的超链接关系进行评级的网页排名算法。该算法基于“从许多优质的网页链接过来的网页,必定还是优质网页”的回归关系,来判定网页的重要性。该算法认为一个网页A对于另一个网页B的连接可以认为是网页A对网页B进行了一次支持性投票,根据每个网页获得的票数来代表自己的重要性。PageRank同时还考虑各网页本身的特性,越重要的页面对支持的页面将给予较高的PageRank值。PageRank算法应用需要满足:①网页之间必须是强连通的;②网页链接不能指向自身。

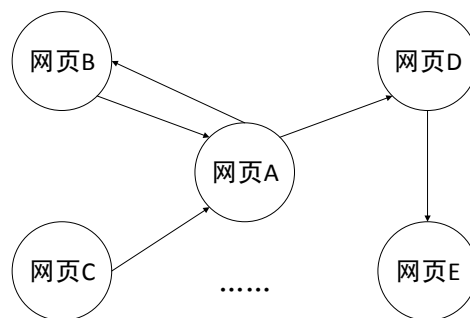


图1 PageRank网页链接示意图

如图1所示为网页之间通过超链接关系形成的链接网络,网页B和网页C链接到网页A,表示网页B和网页C对网页A支持. PageRank值(PR值)的计算公式如下:

$$PR(A) = (1-d) + d \times \sum_i \frac{PR_i}{C(T_i)} \quad (1)$$

公式(1)中, PR(A)为网页A的PageRank值,即网页重要性评分. i 表示链接到网页A的网页,在图1中有网页B和网页C. $C(T_i)$ 表示网页*i*链接到其他网页的数量,当一个网页的链出页面越多,支持力越低. d 表示网页转移概率,即用户关闭当前页面随机浏览另一个页面的概率.

Wang^[12]于2011年的研究中对中文微博和意见领袖特征进行分析,得出关注用户数量、粉丝数量、是否被验证身份和发布的微博数量这四个方面是意见领袖识别的关键.在文献11中提出了微博客用户重要性评分模型(原始模型):

$$Score(U) = A(Focus_U) \times B(T_U) \times Fans_U \times (1 + \alpha V_U) \quad (2)$$

公式(2)中4个乘法因子分别对应于关注用户数量、发布的微博数量、粉丝数量和是否被验证身份,并进行规范化.

Lu^[9]在2015年的研究中对原始模型进行修正,在研究中表明用户之间的转发关系不可忽略,因为粉丝数少的用户的某条微博也可能受到很多关注,同时得到大量的转发,在某一话题下,该用户的影响力也是可以很高的.针对转发关系, Lu在研究中结合了PageRank算法的基本思想,根据用户的转发行为构建了一个关系网络,并考虑了影响力的时间衰减.在文献[8]中提出了意见领袖识别模型(IOLM):

$$IScore(u) = Sp(u) \times Ratio(u, Ctime - LastTime_u) \quad (3)$$

$$Sp(T_i) = (1-\gamma)S(T_i) + \gamma \sum_{i=1}^n \frac{S(T_i)}{C(T_i)} \quad (4)$$

$$Ratio(u, day) = \begin{cases} \frac{1 + S(u) \times (\sum_{i=1}^n S(u))^{-1}}{\log_2 day} & day > 4.59 \\ 1 & others \end{cases} \quad (5)$$

公式(4)中, $S(T_i)$ 表示主题T下用户*t*的影响力,与公式(2)的计算方法相同. $C(T_i)$ 表示用户*i*转载微博的总量,参数 γ 作为权重参数.公式(5)中, day 表示未发表微博的天数.

以上学者提出的三种意见领袖的检测算法仍然存在各自的缺点,本文结合三种算法的优点,以特定主

题下的微博转发关系构建信息传播网络,提出一种在特定话题转发网络下的意见领袖检测方法.

4 特定话题下的意见领袖识别模型

基于Twitter的研究表明:在信息传播过程中,用户影响力与其粉丝数量呈弱相关^[19],粉丝数量只是意见领袖的必要条件.而草根用户在特定话题下的传播中更容易获得用户的关注而提升影响力^[6].由于微博中具有许多社会名人、网络红人等,他们本身具有大量的粉丝,使得他们发布的信息更容易传播出去,但在某些话题下他们更加倾向于对事件进行传播而不具有权威性.传统意见领袖检测方法往往将影响力和传播力等价关联,这样导致粉丝数较少的草根用户的影响力被社会名人等粉丝大户淹没.HITS-BOWR算法^[10]中明确地计算了用户的权威值和中心值,但是该算法仅仅考虑了转发关系,忽略了用户本身发布的微博内容产生的其他影响,经过多次迭代后仍然容易使同一用户的权威值和中心值较为接近.

Liu在2011年研究中提出,意见领袖是主题依赖的^[11],因此用户的影响力往往随话题的领域性而发生变化权威性变化.同样地,传统意见领袖检测方法缺少对主题进行区分,由于不同主题的影响力是相互独立的,用户过去的发布的热门话题微博影响力不能或者只能部分影响到新发布的微博.因此在考虑用户当前话题影响力时,以往发布话题微博的影响力只能作为用户活跃度的一部分.

在话题微博生存期方面,意见领袖的话题影响力并不是简单地随时间的推移进行衰减,因为微博主题本身具有一定的生存期,在主题的热度也随时间而下降,但是早期的意见领袖对该主题造成的影响不并随着时间而下降,相反地,有可能因为粉丝数的转发关系而是影响力继续上升.因此本文针对传统意见领袖存在的以上三个方面的缺点进行改进,使得算法能更加较有效地检测出具有主题相关性和话题权威性的意见领袖.

4.1 用户影响力计算方法

本文将权威性的评分指标分为用户本身在该话题下具有的权威性和其他用户通过转发关系贡献的支持力.特定话题的微博消息转发网络中,用户具有的权威性衡量指标分为:①参与用户发布特定话题的微博的用户数(取代用户粉丝数);②用户发布的微博在消

息网络传播中的节点热度; ③其他用户对该话题微博的专注度. 转发网络中, 转发用户对某一微博的支持力主要体现在当日内该微博得到用户的重视程度.

定义 1. 特定话题下发布的微博的用户具有的权威性评价公式:

$$WScore(u) = A(Forward_u) \times K(u) \times P(TNum_u) \quad (5)$$

公式(5)中, A 函数对用户特定话题下的转发量进行评价, K 函数对话题在信息传播网中的推动力进行评价, P 函数对用户的话题专注度进行评价, 在数据预处理阶段对同一属性的数据进行最大最小值规范化处理. 并且使三个函数的最大值控制在 2 以下.

定义 2. 特定话题下用户发布的微博转发量评价函数

$$A(Forward_u) = \frac{Forward_u - \frac{1}{n} \sum_{l \in Top(n)} Forward_l}{Max(Forward_u, \frac{1}{n} \sum_{l \in Top(n)} Forward_l)} + 1 \quad (6)$$

公式(6)中, Forward_u 表示用户 u 在特定话题 i 下发表的微博得到的转发量. n 表示意见领袖参考数量, 由于微博转发数量呈现长尾效应(在下一节实验部分证明), 极大部分微博只有很少的转发量, 因此传播网络中的总体转发量均值很低, 只需要研究转发量排名前 n 个的用户, 在研究中 n 取 60. Top(n) 表示转发量前 n 名的用户的标签. Max() 为取最大值函数. 转发量可以较好地表征用户对微博的兴趣程度, 因此该公式可以评价出微博受到的关注程度. 由于转发量的长尾现象, 导致大部分用户的值为负数, 因此在这里最小值都设为 0.001.

信息传播网中的时间节点热度评价函数 $K(time_u) = 2 \times k_u$, 值域为 [-2, 2]. 如公式(7), α 取值于该节点在话题 i 的每日转发量曲线图中的斜率, 表示该话题的热度在该时间节点位置处于上升阶段还是下降阶段. 上升阶段表示该节点发布的微博对于话题具有一定的推动作用, 用户在该阶段更容易受到该话题的影响. 这里实际上取代了公式(4)中用户影响力时间衰减因素, 而是根据话题的推动作用作为该时段影响力变化的评价. 特定话题的用户转发量以散点图的形式体现, 信息统计窗口以 2 小时为一个单位. 拟合时间段影响力变化曲线在某点切线的斜率可以用最小二乘法来计算, 如图 2.

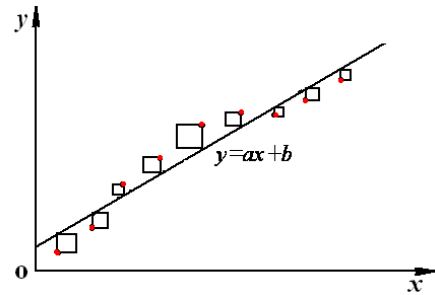


图 2 最小二乘法计算拟合曲线斜率

$$\alpha = \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i - m \sum_{i=1}^n (x_i y_i)}{(\sum_{i=1}^n x_i)^2 - m \sum_{i=1}^n x_i^2} \quad (7-1)$$

$$k_u = \frac{\arctan(\alpha_u)}{\pi} \quad (7-2)$$

定义 3. 用户对话题的专注度

$$P(TNum_u) = \sqrt{\frac{1}{\ln(1 + TNum_u^2)}} \quad (8)$$

公式(8)中, TNum_u 表示用户在发布特定话题当日的转发其他话题微博的数量.

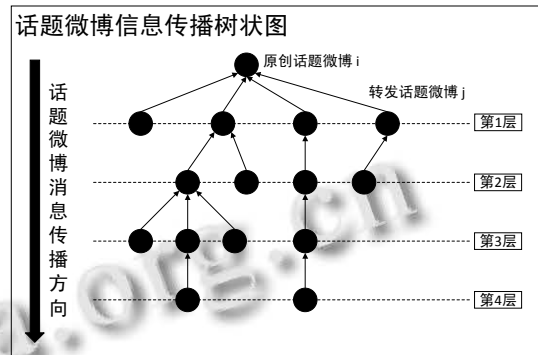


图 3 话题微博信息传播树状图

在微博消息传播网络中, 用户通过转发关系对微博信息进行扩展, 同时微博与转发微博之间满足一对多关系, 因此特定话题微博的消息传播网络可以以树状图的形式展示. 如图 3 所示, 在信息传播的树状图中有明显的层次关系, 第 i 层的微博与第 i+1 层的微博有一对多的转发关系, 与 i-1 层有一对一的转发关系, 相邻两层之间有直接的转发关系. 并且由于影响力的支持直接影响上一层微博, 相隔层数越多, 支持力越弱, 因此可以通过迭代关系进行计算支持力.

定义 4. 特定主题的意见领袖识别模型(TOLM):

$$TScore(u) = \gamma WScore(u) + (1 - \gamma) \sum_{l(l,u) \in E} WScore(l) \quad (9)$$

公式(9)中, TScore(u)即为特定主题的意见领袖识别模型中用户U的影响力评分. γ 为权重参数, 实验中取0.8, 1表示具有通过转发关系相连接的用户, 通过迭代关系可知, 相邻层数越多的用户, 具有的权重越低.

4.2 数据收集和模型检测框架

研究中所分析的数据都来自于新浪微博中热点话题下用户的信息转发网络. 当前流行的数据获取方法主要有两种: 1)由新浪微博开放平台提供的应用程序编程接口(简称微博 API); 2)网络爬虫程序实时抓取. 一方面, 由于微博 API 对接口开放的限制, 部分接口为高级权限接口, 需要申请才可以调用, 不利于数据收集的全面性; 另一方面, 由于微博 API 数据搜索返回的数量限制, 不利于数据收集的完整性. 因此在研究中主要还是通过设计爬虫脚本来抓取微博数据.

研究中的模型框架如图 4 所示, 第一步中需要人工输入需要检测的话题关键词信息, 或者通过词共现聚类方法找出特定话题下的热点词汇, 关键词信息直接关系着后续微博的话题类别. 微博信息分析阶段, 主要是提取三部分内容: ①用户(近期微博发布情况); ②话题微博(转发数, 转发用户, 时间); ③粉丝(参与转发的用户信息). 在数据库信息反馈阶段, 由于话题的舆论走向和意见领袖的影响力与日发生变化, 需要定时更新话题热点词汇以及用户的影响力. 在同一话题下, 原创微博之间具有相互独立性, 并且都以树状网络的传播方式呈现; 总体上相同话题下的多个原创微博信息传播呈现出森林结构.

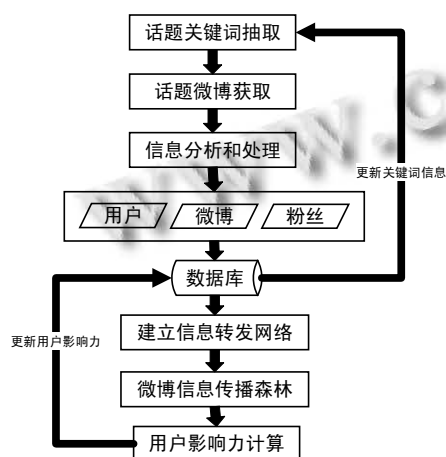


图 4 模型检测框架

5 实验结果分析

本实验通过网络爬虫软件, 于 2016 年 3 月 2 日至

3 月 18 日对新浪微博热门话题下的微博信息进行抓取, 并存入数据库. 如表 1 所示, 在研究中提取了三个话题类别, 并剔除重复微博和干扰用户(包括广告用户、话题无关用户、僵尸用户等).

表 1 研究中的话题微博内容

序号	话题关键词	总数据	有效数据	话题提取日期
1	人机围棋大战	4137	1828	2016 年 3 月 4 日至 3 月 18 日
2	张怡宁 福原爱	2772	1104	2016 年 3 月 6 日至 3 月 19 日
3	唐安琪 冈本白贤	3322	1588	2016 年 3 月 2 日至 3 月 14 日

实验一. 微博话题生存期变化情况

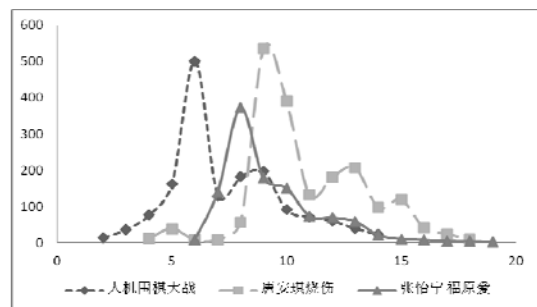


图 5 话题微博生存期数量变化示意图

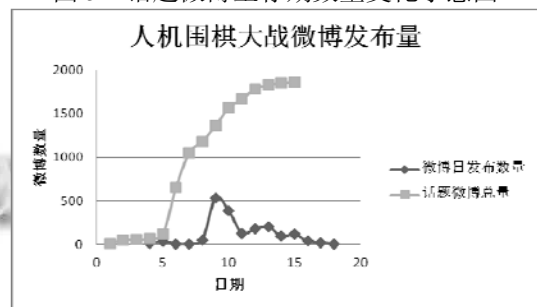


图 6 两周内话题微博日发布量和总量变化示意图

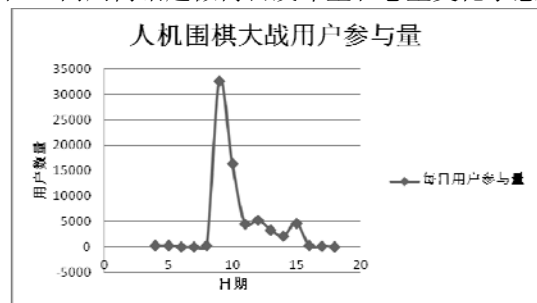


图 7 两周内用户日参与量变化示意图

由图 5 中“人机围棋大战”、“唐安琪烧伤”和“张怡

宁 福原爱”三个话题下的微博可以看出，热门话题下的微博发布情况，遵循从潜伏期到爆发期，再从爆发期到冷淡期这一规律，并且潜伏期到爆发期的过程较短。由图 6 和图 7 可以看出，在话题的爆发期间内，用户的参与度达到顶峰，并且微博发布量也随之达到顶峰，最后微博发布总量趋向于平缓，也表示着话题的生存期结束。因此可以认为，意见领袖的检测关键时间就在潜伏期到爆发期这一阶段，这一阶段微博发布量大，用户参与度高；而爆发期到冷淡期，随着用户参与量的降低，话题热度的下降，用户影响力难以扩散。因此，引证了意见领袖的出现具有时效性这一特点。同样，话题潜伏期到爆发期这一阶段的微博发布量变化曲线的斜率突增，有利于对事件的发展进行判断。

实验二. 转发网络中的长尾现象

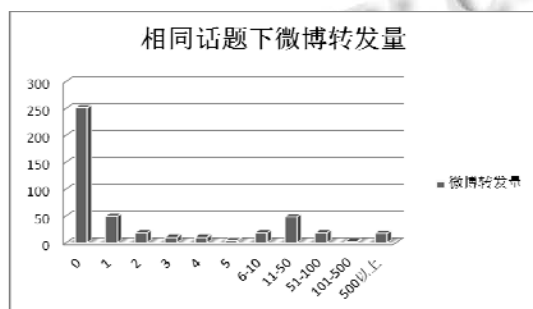


图 8 用户相同话题下的微博转发量呈长尾现象

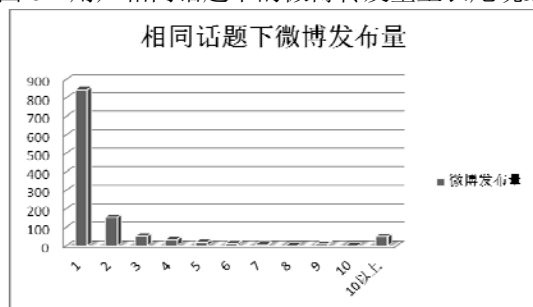


图 9 用户在相同话题下的微博发布量呈长尾现象

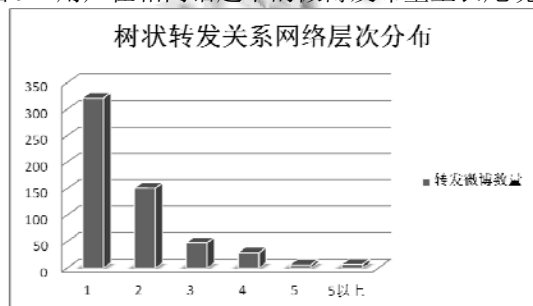


图 10 转发关系树状图中转发层次的微博数量呈长尾现象

用户对于某话题的关注行为，主要体现在发布与话题相关的微博，或者转发话题相关的微博。在图 8 和图 9 中我们可以看出，用户们普遍对某话题的关注行为不会超过两次，因此在用户影响力评估时，不需要像传统意见检测那样对关注行为的数量耗费太多资源。在话题“人机围棋大战”的数据库中，微博转发量超过 100 的原创微博数量共有 60 篇，研究中对这 60 篇微博分别建立转发关系树状，对转发层次的平均数进行分析，同样发现树状图中的转发层次与微博数量呈长尾现象，如图 10。由于用户转发关系共现的支持力以层次关系迭代运算，为了节约成本提高实时性方面，可以对迭代次数进行控制，一般控制在 5 以内。

实验三. 意见领袖的排名

表 2 不同指标下的意见领袖排名

用户自身权威值	权威值+支持力	转发量	粉丝量
马伯庸	央视新闻	央视新闻	头条新闻
央视新闻	头条新闻	深圳供电	环球资讯广播
头条新闻	马伯庸	头条新闻	央视新闻
谭文胜微博	深圳供电	新浪体育	中国新闻周刊
新浪科技	新浪体育	马伯庸	每日经济新闻
新浪体育	谭文胜微博	央视新闻	新闻晨报
央视新闻	新浪科技	杜蕾斯官方微博	中国新闻网
深圳供电	央视新闻	谭文胜微博	新快报
杜蕾斯官方微博	杜蕾斯官方微博	新浪科技	财经网
摘星手 010	360 手机	360 手机	经济观察报

在表 2 列出了不同指标下的意见领袖排名情况，可以用用户自身权威值的评价可以很好地突出草根用户，转发量的评价可以表示出用户对其他用户的影响力，而粉丝量与用户的影响力只是呈现弱相关关系，因此用户自身权威值加支持力来评价一个用户的影响力能更加贴近实际情况。在表 3 中列出了前 12 名意见领袖的详细信息，可以发现在 3 月 9 日这一天的意见领袖数量较多，因为该话题在这一天的消息传播最为迅速，话题的推动力较强，因此用户的影响力更容易扩散，同时通过粉丝数和转发量排名的比较，更能模型中通过话题参与用户来取代粉丝量的有效性。

表3 模型评分前12名意见领袖信息

模型评分排名	评分	排名	粉丝数	排名	转发量	排名	发布日期
央视新闻	1.7389	1	42129029	4	6541	1	3月9日
头条新闻	0.8132	2	49689957	1	2172	3	3月9日
马伯庸	0.7389	3	3133790	181	1409	5	3月9日
深圳供电	0.6910	4	402995	323	2182	2	3月15日
新浪体育	0.5862	5	8530212	71	1552	4	3月9日
谭文胜微博	0.5291	6	3032837	183	1062	7	3月10日
新浪科技	0.5258	7	11690597	45	1055	8	3月9日
杜蕾斯官方微博	0.4107	8	1533375	234	1218	6	3月10日
NASA 中文	0.3045	9	2779845	192	741	10	3月9日
360 手机	0.2760	10	1108219	265	842	9	3月12日
摘星手 010	0.1961	11	154809	360	621	12	3月11日
搜狗号码通	0.1946	12	576754	300	693	11	3月10日

6 结语

研究中提出的基于 PageRank 的意见领袖检测方法主要以 Wang 和 Lu 提出的基本模型为基础,通过转发关系建立了信息传播网络,结合 PageRank 的影响力评价方法来对特定话题下的意见领袖进行识别,提出了特定主题的意见领袖识别模型(TOLM)。未来工作中将进一步结合话题微博下的情感信息这一属性来评价意见领袖的影响,同时加强话题信息传播网络的完整性。

参考文献

- 1 中国互联网络信息中心.第34次中国互联网络发展状况统计报告.互联网天地,2014,(7).
- 2 Lazarsfeld PF, Berelson B, Gaudet H. The people's choice. Eco-Architecture: Harmonisation between Architecture and Nature, 1944, 18(Jan): 154.
- 3 Rogers EM, Shoemaker FF. Communication of innovations; a cross-cultural approach. Man, 1971, 9(2): 476.
- 4 Brin BS, Page L. The anatomy of a large scale hypertextual Web search engine. Computer Networks and ISDN Systems, 2012.
- 5 Jon M, Kleinberg. Authoritative sources in a hyperlinked environment. Journal of the Acm, 1998, 46(5): 604-632.
- 6 Weng J, Lim EP, Jiang J, et al. TwitterRank: Finding topic-sensitive influential twitterers. Proc. of the Third International Conference on Web Search and Web Data Mining, WSDM 2010. New York, NY, USA. February 4-6, 2010. 261-270.
- 7 肖宇,许炜,商召玺.微博用户区域影响力识别算法及分析. 计算机科学,2012,39(9):38-42.
- 8 原福永,冯静,符茜落.微博用户的影响力指数模型.现代图书情报技术,2012,(6):60-64.
- 9 卢伟胜,郭躬德.基于特定话题的微博意见领袖在线检测方法.计算机应用与软件,2015,32(5):70-74.
- 10 熊涛,何跃.微博转发网络中意见领袖的识别与分析.现代图书情报技术,2013,(6):55-62.
- 11 刘志明,刘鲁.微博网络舆情中的意见领袖识别及分析.系统工程,2011,(6):8-16.
- 12 王君泽,王雅蕾,禹航,等.微博客意见领袖识别模型研究.新闻与传播研究,2011,(6):81-88.
- 13 李玉贞,胡勇,熊熙,等.微博意见领袖的评估模型.信息安全与通信保密,2013,(2):79-81.
- 14 丁兆云,周斌,贾焰,等.微博中基于多关系网络的话题层次影响力分析.计算机研究与发展,2013,50(10):2155-2175.
- 15 王晨旭,管晓宏,秦涛,等.微博消息传播中意见领袖影响力建模研究.软件学报,2015,26(6):1473-1485.
- 16 赵静.影响力扩散概率模型及其用于意见领袖发现研究[硕士学位论文].重庆:重庆邮电大学,2013.
- 17 周而重,钟宁,黄佳进.基于意见领袖引导作用的网络舆论演化研究.计算机科学,2013,40(11):287-290.
- 18 张腊梅,黄威靖,陈薇,等.EMTM:微博中与主题相关的专家挖掘方法.计算机研究与发展,2015,52(11):2517-2526.
- 19 Asur S, Huberman BA, Szabo G, et al. Trends in social media: Persistence and decay. Ssrn Electronic Journal, 2011.