

## 基于社会网络中心性分析的微博信息传播研究

——以 Sina 微博为例

Research on Microblog Information Dissemination Based on SNA Centrality Analysis

——A Case Study with Sina Microblog

平 亮 宗利永

(上海理工大学管理学院,上海 200093)

[摘要] 基于社会网络理论,结合微博用户之间的“关注”与“被关注”信息传播的网络拓扑关系,利用各种中心度和中心势测度指数,分别从点度中心性、中间中心性和接近中心性三个方面对微博社会网络的中心性进行了分析,最后提出了相应的启示。

[关键词] 微博 社会网络分析 中心性 权力

[中图分类号] G350 [文献标识码] A [文章编号] 1003-2797(2010)06-0092-06

[Abstract] This paper adopted the methods of centrality and central potential indexes. Combined with the micro blog users network topology of “concerning” and “concerned” and analyzed the centrality of the microblog network from point centrality, betweenness centrality and closeness centrality finally the conclusion and implications are put forward.

[Key words] MicroBlog Social network analysis Centrality Power

“中心性”是社会网络分析的重点之一,个人或组织在社会网络中具有怎么样的权力,或者说居于怎样的中心地位,对于信息在整个网络中如何传播,及其传播效果如何都有着十分重要的意义。本文从社会网络的研究视角出发,通过观察一组微博用户,利用 UCINET 软件测量网络结构的各种中心性,了解每个节点的中心度情况,以期从“中心性”的描述及其测量方法出发对微博信息传播进行分析。

## 1 社会网络视角下的微博信息传播

微博,即微博客(MicroBlog),是一个基于用户关系的信息分享、传播以及获取平台,用户可以通过各种客户端以 140 字左右的文字更新信息,相对于强调版面布置的博客来说,微博的内容简单,在语言的编排组

织上对用户的要求较低。由于用户可以将自己的微博与手机绑定,用户更是可以随时随地地更新信息,不再受时间与空间的限制,使得信息传播更加及时,特别是对突发事件的报道和反馈。

对于传统博客来说,用户的关注属于一种“被动”的关注状态,写出来的内容其传播受众并不确定;而微博的关注则更为主动,只要轻点“follow”(在新浪微博中表现为“关注”)即表示你愿意接受某位用户的即时更新信息,“由于这种半广播半实时交互的微博客机制,使得用户组成多个交流分享的小圈子,群体传播在这里得到凸显”<sup>[注1]</sup>。同时,对于普通人来说,微博的关注友人大多来自事实的生活圈子,用户的一言一行不但起到抒发感情,记录日常生活的作用,更重要的是

[基金项目] 本文系上海市研究生创新基金项目“互联网危机信息传播及舆情演变模型”(JWCXSL1001)研究成果之一。

[作者简介] 平亮,男,1982年生,硕士研究生;通讯作者:宗利永,男,1981年生,博士研究生。

维护了人际关系。

微博信息传播是基于一定的社会网络基础的传播模式。特别是上述第三点提到的“人际圈的影响”,显然在网络上人与人的联系不是随机的,毫无规律可寻的。人们由于教育背景、兴趣爱好、职业背景等方面的差异导致其对信息的偏好也不尽相同,人们会根据自己的爱好,而有选择地对其他用户进行关注,从而形成一个虚拟“圈子”。在这个“圈子”中,人与人不是孤立的,之间是有联系的,故人与人之间的关系可以抽象成一个社会网络图中的节点与节点之间的连线,“节点”可视为信息传播主体,“关系”可用于表示信息传播主体之间某些特定的联系,“边”表示信息传播的路径。不难看出,微博用户及其之间的关系本质上是一个社会网络。

## 2 社会网络中心性分析的三种方法

在社会网络“中心性”的描述中,中心度与中心势是两种重要的测量方法。中心度指的是一个点在网络中居于核心地位的程度,而中心势考察的是整个图的整体整合度或者一致性,也就是一个图的中心度。而社会网络的中心性又可分为点度中心性、中间中心性、接近中心性三种。其中每一种中心性都有中心度和中心势两种指数来描述。其具体分类如表1:

表1

表示方法	中心性	中心度	中心势
点度中心性	点度中心度	绝对中心度	图的点度中心势
		标准化中心度	
中间中心性	中间中心度	绝对中心度	图的中间中心势
		标准化中心度	
接近中心性	接近中心度	绝对中心度	图的接近中心势
		标准化中心度	

(1) 点度中心性(point centrality)。点度中心度指的是在社会网络中,一个行动者(actor)与很多其他行动者有直接联系,该行动者就处于中心地位,从而拥有较大的权力。点度中心势:中心势特指作为一个图的中心度,中心势指的不是点的相对重要性而是图的整体整合度或者一致性。在中心性的分析中,有时候我们关注的不是点,而是整个图,研究不同图是否有不同的中心趋势。

(2) 中间中心性(betweenness centrality)。中间中心度:如果一个行动者处于许多交往网络的路径上,可以认为此人居于重要的地位。因为他具有控制其他两人之间交往的能力。中间中心度测量的是行动者对资源信息控制的程度,如果一个点处于许多其他点对的测地线(最短途径)上,就说该点具有较高的中间中心度。

(3) 接近中心性(整体中心性)(closeness centrality)。接近中心度:点度中心度仅仅是点中心度的一种测量,有时候我们更关注一个行动者与网络中其他行动者的接近程度。这便引出了点的接近中心度(closeness centrality):一种不受他人控制的测度。当行动者越是离其他人接近,则越是在信息传播中不依赖他人。因我们称此人有较高的中心度。因为一个非核心位置的成员必须“通过他人才能传递到信息”,故如果一个点与网络中其他各点的距离都很短,则该点是整体中心点。

在实际测量“中心性”的时候,对测度方法的选择,弗里曼认为“这依赖于研究问题的背景,如果关注交往活动,可采用以度数为基础的测度;如果研究对交往或者对信息的控制,可利用中间中心度;如果研

表2 点的中心度和图的中心势表达<sup>[注2]</sup>

	点度中心性	中间中心性	接近中心性
绝对点度中心度	$C_{ADi} = i \text{ 的度数}$	$C_{ABi} = \sum_j \sum_k b_{jk}(i) \quad j \neq k \neq i$ $j < k$	$C_{APi}^{-1} = \sum_{j=1}^n d_{ij}$
标准化中心度	$C_{AD}(i) / (n-1)$	$C_{RBi} = \frac{2C_{ABi}}{n^2 - 3n + 2}$	$C_{RPI}^{-1} = \frac{C_{APi}^{-1}}{n-1}$
图的中心势	$C_{RD} = \frac{\sum_{i=1}^n (C_{RDmax} - C_{RDi})}{n-2}$	$C_B = \frac{\sum_{i=1}^n (C_{Rmax} - C_{RBi})}{n-1}$	$C_c = \frac{\sum_{i=1}^n (C_{RCmax} - C_{RCi})}{(n-2)(n-1)}$

究信息传递的独立性和有效性,可采用接近中心度(Freeman,1979:226)。

3 微博社会网络中心性的实证分析<sup>[注3]</sup>

在新浪微博中,用户大致可以分为“名人”和“草根”两部分,相对而言“名人”受到更多的“关注”,故在信息及其“网络舆情”的传播中,起到更为重要的作用。本文列举的微博用户,都是有10万以上的人数“关注”的“名人”。本文在取样时采取“滚雪球”的方法,即在随机确定一个“名人”微博用户后,观察其“关注”对象,并将被“关注”人数超过10万的用户记录下来,之后再记录下来的用户采取同样的方法观察记录,从而得到以下一组用户的信息<sup>[注4]</sup>。由于新浪微博的用户群十分庞大(数百万用户),以下一组用户仅是整个微博社会网络的一个子网络,嵌套在整个网络之中,但由于选取样本属于用户中的“名人”,具有相当程度的辐射广度。本文通过对这个小型的社会网络进行描述,从而对信息或者“网络舆情”在整个微博社会网络中传播进行探讨。

3.1 微博社会网络的“关注矩阵”及其关联图

经过观察记录,将用户之间“关注”与被“关注”的关系用一个邻接矩阵的形式表达出来。在这个方阵之中行表示“关注”者,列表示被“关注”者。“1”表示“关注”这种关系存在,“0”则表示不存在。我们可以称这个矩阵为“关注矩阵”,见下表:

用户	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	—	0	0	0	0	0	0	0	0	0	0	0	0	1
2	0	—	0	0	0	0	0	0	0	0	0	1	0	0
3	0	0	—	0	0	0	1	0	0	0	0	0	0	0
4	0	0	0	—	1	0	0	0	0	0	0	0	0	0
5	0	0	1	1	—	1	1	1	1	1	1	1	1	1
6	1	0	1	1	1	—	1	1	1	1	1	1	0	1
7	0	0	0	0	1	1	—	1	0	0	0	1	1	1
8	1	1	1	1	1	1	1	—	1	1	1	1	0	1
9	0	0	0	1	0	0	1	0	—	1	1	1	1	0
10	0	0	0	0	1	0	0	0	1	—	0	1	0	0
11	0	0	0	0	0	0	0	0	1	0	—	0	0	0
12	0	1	0	0	1	0	0	1	0	1	0	—	0	0
13	0	0	0	0	1	0	0	0	0	0	1	1	—	0
14	0	0	0	0	1	0	1	1	1	1	0	1	1	—

通过 UCINET 软件的画图功能可以将网络关系图直观地表现出来。见图1。

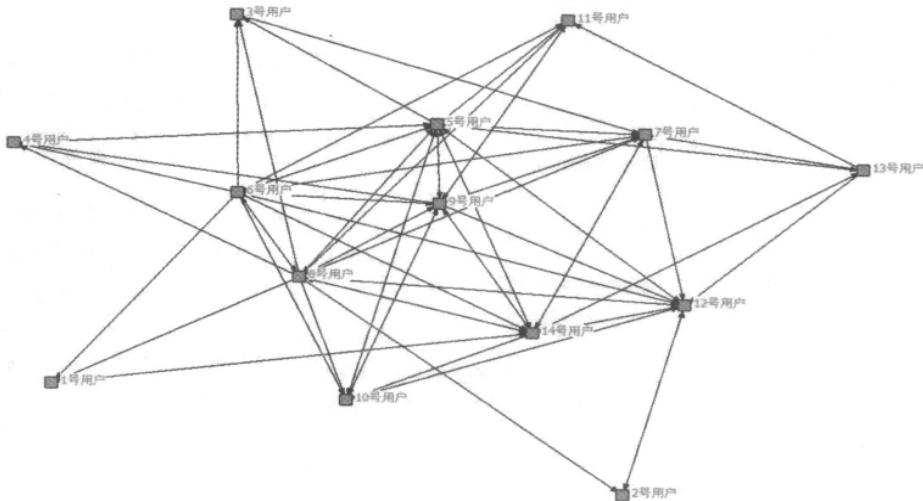


图1 微博社会网络的关联图

根据“关注矩阵”。故可以判断这一组用户构成的社会网络关系图为有向图。在这个有向关系图中,

每一个结点表示一个用户,他们之间的连线表示存在这种“关注”与被“关注”的关系。箭头从结点A指向

结点 B 表示 A “关注” B。一个结点指向其他结点的箭头的数量记做点出度 (outdegree), 一个结点被其他结点所指的箭头的数量记做点入度 (indegree)。

由于这些结点之间存在着这种有向关系, 即如果 A “关注” B 时, B 发布的消息, 很快会在 A 的主页上显示, 这样信息从 B 传播到 A。故在微博社会网络之中, 信息的传播方向是与 “关注” 的方向相反的。

### 3.2 各种中心度与中心势的测算

本文结合微博用户之间的 “关注” 与 “被关注” 关系, 通过分析各种中心度和中心势指数, 找出那些用户在这个社会网络中处于核心地位, 即找出哪些用户在微博信息传播的过程中的 “权力” 更大, 能够在较大程度上影响整个社会网络上的信息传播。

#### (1) 点度中心性分析

将上述的 “关注矩阵” 输入 UCINET 软件, 计算其点度中心度及其中心势, 结果如下:

#### FREEMANS DEGREE CENTRALITY MEASURES

Diagonal valid? NO

Model: ASYMMETRIC

Input dataset: D:\Ucinet 6\DataFiles\关注矩阵

		1	2	3	4
		OutDegree	InDegree	NrmOutDeg	NrmInDeg
		-----	-----	-----	-----
8	8 号用户	12.000	5.000	92.308	38.462
5	5 号用户	11.000	8.000	84.615	61.538
6	6 号用户	11.000	3.000	84.615	23.077
14	14 号用户	7.000	6.000	53.846	46.154
9	9 号用户	6.000	6.000	46.154	46.154
7	7 号用户	6.000	6.000	46.154	46.154
12	12 号用户	4.000	9.000	30.769	69.231
13	13 号用户	3.000	3.000	23.077	23.077
10	10 号用户	3.000	6.000	23.077	46.154
4	4 号用户	1.000	4.000	7.692	30.769
1	1 号用户	1.000	2.000	7.692	15.385
2	2 号用户	1.000	2.000	7.692	15.385
3	3 号用户	1.000	3.000	7.692	23.077
11	11 号用户	1.000	5.000	7.692	38.462

#### DESCRIPTIVE STATISTICS

		1	2	3	4
		OutDegree	InDegree	NrmOutDeg	NrmInDeg
		-----	-----	-----	-----
1	Mean	4.857	4.857	37.363	37.363
2	Std Dev	3.925	2.065	30.195	15.887
3	Sum	68.000	68.000	523.077	523.077
4	ariance	15.408	4.265	911.726	252.385
5	SSQ	546.000	390.000	32307.693	23076.924
6	MCSSQ	215.714	59.714	12764.159	3533.389
7	Euc Norm	23.367	19.748	179.743	151.911
8	Minimum	1.000	2.000	7.692	15.385
9	Maximum	12.000	9.000	92.308	69.231

Network Centralization ( Outdegree ) = 59.172%

Network Centralization ( Indegree ) = 34.320%

Actor-by-centrality matrix saved as dataset  
FreemanDegree

从分析结果来看, 不同的用户表现出不同的点入度和点出度。点入度的含义是关系 “进入” 的程度, 在这里一个用户表示被其他用户 “关注” 的程度。点出度表示一个用户 “关注” 其他用户的程度。从结果中可以看出, 点入度比较高的用户, 即更受人 “关注” 的用户为 12 号用户 (点入度为 9), 5 号用户 (点入度为 8)。说明他们在这个网络上信息传播的过程中拥有较大的权力, 他们发布的消息为更多人们所注意。

整个网络的标准化点入度中心势和点出度中心势分别为: 34.320% 和 59.172%。中心势越接近 1, 说明网络越具有集中趋势 (centralization)。可以看出 “关注” 的中心势更大一些, 说明 “关注” 他人的用户更具集中趋势。同样, 被 “关注” 中心势也达到了 34.320%, 也有着比较明显的集中趋势, 也就是说被 “关注” 的用户有着明显的集中趋势。

#### (2) 中间中心性分析

将上述 “关注” 矩阵输入 UCINET 软件, 计算中间中心性度及其中心势, 结果如下:

## 基于社会网络中心性分析的微博信息传播研究——以 sina 微博为例

Analysis of Microblog Information Dissemination Based on SNA Centrality Modeling: A Case Study with Sina Microblog

平 亮 宗利永

## FREEMAN BETWEENNESS CENTRALITY

1		2	
Betweenness		nBetweenness	
-----	-----	-----	-----
5	5 号用户	41.011	26.289
12	12 号用户	24.078	15.434
8	8 号用户	23.900	15.321
7	7 号用户	19.067	12.222
14	14 号用户	16.578	10.627
9	9 号用户	15.167	9.722
6	6 号用户	5.789	3.711
10	10 号用户	1.289	0.826
13	13 号用户	1.000	0.641
4	4 号用户	0.622	0.399
11	11 号用户	0.500	0.321
2	2 号用户	0.000	0.000
3	3 号用户	0.000	0.000
1	1 号用户	0.000	0.000

Network Centralization Index = 20.96%

点的中间中心度测量的是一个点 C 在整个网络中对信息的流动或者是传播的控制作用的大小,即信息要想从结点 A 传达到结点 B 在多大程度上要依赖于结点 C。我们可以从分析结果中清晰地看到,5, 12, 8, 7, 14 和 9 号用户的中间中心度是比较高的。也就是说其他各个用户想获得消息对上述几个用户的依赖程度是比较高的,故这几个用户在这个网络上的权力比较大,能够在较大程度上控制信息的流动。而整个网络的中间中心势为 0.2096,并不是很高,就整个网络而言,中间中心度并不是特别大,即在整个网络中大部分的结点不需要别的节点作为桥接点,便可以得到信息。

## (3) 接近中心性的测量

将上述“关注矩阵”输入 UCINET 软件,计算接近中心性及其中心势,结果如下:

## CLOSENESS CENTRALITY

## Closeness Centrality Measures

		1	2	3	4
		inFarness	outFarness	inCloseness	outCloseness
-----	-----	-----	-----	-----	-----
12	12 号用户	17.000	22.000	76.471	59.091
5	5 号用户	19.000	15.000	68.421	86.667
7	7 号用户	21.000	20.000	61.905	65.000
14	14 号用户	21.000	19.000	61.905	68.421
10	10 号用户	21.000	24.000	61.905	54.167
8	8 号用户	22.000	14.000	59.091	92.857
9	9 号用户	22.000	22.000	59.091	59.091
11	11 号用户	24.000	34.000	54.167	38.235
13	13 号用户	25.000	24.000	52.000	54.167
4	4 号用户	25.000	27.000	52.000	48.148
6	6 号用户	26.000	15.000	50.000	86.667
2	2 号用户	28.000	34.000	46.429	38.235
3	3 号用户	28.000	31.000	46.429	41.935
1	1 号用户	32.000	30.000	40.625	43.333

Network in-Centralization = 44.90%

Network out-Centralization = 74.36%

接近中心度测量的是一个行动者不受他人控制的程度,与上述的两个中心度相反,该值越小,说明该点越处于核心位置,因为根据接近中心度的含义,当该值越小时,说明该点与其他点的距离和越小,说明该点距离其他各点越近,在获取信息时越不容易受其他点的控制。从分析结果中可以看到,越靠前的点的中心度越高。如 12 号用户,他发布的信息传递到其他所有点的距离之和只有 17。获取其他所有点的信息则要相对困难,因为距离和为 22。从发布信息到其他各点的方便易难程度而言,所有用户的排序为 12, 5, 7, 14, 10, 8, 9, 11, 13, 4, 6, 2, 3, 1, 越靠前则越不容易受到他人控制,越独立。而获取信息的易难程度而言,排序为 8, 5(6), 14, 7, 12(9), 10(13), 4, 1, 3, 2(11)。靠前的容易获取信息,不易受人控制,独立性强。

不难看出此网络的接近中心势是相当高的。Network in-Centralization 和 Network out-Centralization 分别为 0.449 和 0.7436。说明在网络上每个点发布信息时,信息都能比较顺利地到达其他各点,不会受到太多的控制。相比之下,网络上每个点获取信息则更



加容易,因为每个点在相当程度上是独立的,在获取信息是很少受到其他点的控制。

#### 4 结论与启示

(1) 点度中心性描述的是某个体在发布信息或者获取信息时“权力”的大小。当一个微博用户点出度中心度高时,说明他能够更多地获取其他微博用户的信息。在获取信息方面有着比较大的“权力”。同样当一个用户的点入度中心度高时,说明此微博用户被很多其他微博用户“关注”,说明他在发布信息时有较大的“权力”。中间中心性,指的是控制信息流动的“权力”,当一位用户的中间中心度高时,说明他在很多的用户之间起着“桥接点”的作用,故说他在控制其他用户之间信息传递时的“权力”大;接近中心性,指的是使得信息快速便捷地到达各个用户的“权力”。当一个用户的接近中心度高时,说明他与这个社会网络中的任何一个用户的网络距离都很近,他能很便捷,或者说比较直接地(较少受其他用户控制)获得各个用户信息,以及能便捷地将信息传递到其他用户,较少依赖其他用户。

(2) 社会网络的网络结构决定着网络成员的“权力”大小,而中心性便是描述“权力”的一种比较理想方法。“权力”影响着信息的流动,这里的“权力”不是社会网络成员的自身属性,而是网络中行动者之间存在的一种互动模式。社会网络是从“关系”的角度出发定量地界定权力的,并且给出了关于社会权力的具体形式化定义,即各种中心度和中心势指数来表示权力的概念。在一个社会网络中,信息及其资源的传递往往受“权力”的影响。本文通过观察现实中一组微博用户,对各种中心性进行了测算并且得出了结果,判断他们在微博社会网络上的信息传播过程中,每个个体的“权力”大小以及排序,这里的“权力”不是个体或者群体的特征,可以理解为微博信息发布者与接受者之间实际存在的或者潜在的互动模式。

(3) 在微博的每一个“圈子”中,用户由于其自身属性及其所在网络中位置的不同,其在舆情传播过程中的作用也大不相同。在微博这个巨大的社会网络中,每一个用户在网络信息传播中的“权力”是不同的。“权力”大的往往会被称为“意见领袖”。“意见领袖”对信息传播起到比较大的作用,并能在一定程度上引导舆论。新浪微博将用户大概分为“名人”和“草

根”两部分。一部分用户是经过新浪实名认证的“名人”,由于“名人”之前其他媒体形式上的频繁出现,已经成为家喻户晓的人物,故在新浪微博上也会有很多人“关注”(有些“名人”已超过百万的“粉丝”),这些名人在这个网络上的位置注定会使他们成为“意见领袖”。而“意见领袖”由于中心性的不同,“权力”的形式也会有所不同,在微博网络上对信息传播的影响也会存在巨大的差异。

文中使用的实例样本量较小,并不能描述出一个大规模的网络结构,仅试图通过社会网络中心性分析的若干测度方法对微博的信息传播特点加以描述,对不同用户在信息传播过程中的“关注”与“被关注”的关系进行分析,希望能为网络信息传播的研究提供一个新的视角。

#### 注释

- 1 孙卫华,张庆永. 微博客传播形态解析. 传媒观察 2008(10)
- 2 此表描述的是几种测度的具体公式,其中 $n$ 表示结点数,表示点 $i$ 和点 $j$ 之间的测地线(两个节点之间最近的途径)距离。
- 3 本文实证数据来自于2010年3月30日12:00采样的新浪微博(<http://t.sina.com.cn>)。
- 4 微博地址 1 号用户 <http://t.sina.com.cn/tangjun>; 2 号用户 <http://t.sina.com.cn/renzhiqiang>; 3 号用户 <http://t.sina.com.cn/shiyuzhu>; 4 号用户 <http://t.sina.com.cn/wangshi>; 5 号用户 <http://t.sina.com.cn/kaifulee>; 6 号用户 <http://t.sina.com.cn/1497385214>; 7 号用户 <http://t.sina.com.cn/wanglifan>; 8 号用户 <http://t.sina.com.cn/xuxiaoping>; 9 号用户 <http://t.sina.com.cn/honghuang>; 10 号用户 <http://t.sina.com.cn/ruichenggang>; 11 号用户 <http://t.sina.com.cn/luqiuluwei>; 12 号用户 <http://t.sina.com.cn/panshiyi>; 13 号用户 <http://t.sina.com.cn/huangjianxiang>; 14 号用户 <http://t.sina.com.cn/yuminhong>

#### 参考文献

- 1 王夏洁,刘红丽. 基于社会网络理论的知识链分析. 情报杂志, 2007(2)
- 2 李士洋. 社会网络理论、技术发展与组织研究. 现代商业, 2010.
- 3 邱均平,马瑞敏,于长福. 图林博客的社会网络分析. 图书情报工作 2008(11)
- 4 姜胜洪. 试论网上舆情的传播途径、特点及其现状. 社科纵横, 2008(1)
- 5 林聚任. 社会网络分析:理论、方法与应用. 北京:北京师范大学出版社 2009.

(收稿日期:2010-05-13)