

# 去“巴尔干化”： 微博社会“中心化”趋同现象与结构

徐翔 夏敏 唐颖 王晓虹

**【摘要】**在社交媒体用户的连接和趋同过程中,各类用户局部的碎片化、区隔化得以强调,但对碎片化背后全局的同质性和统一性关注不足。重要的问题不仅仅是社交网络是否割裂为“巴尔干化”的一个个孤岛,而是它是否处于一个单一的“整体社会孤岛”之中。本文认为,社交媒体各类用户并非是巴尔干化或去中心化的,而是存在“中心化”的趋同结构,子类的碎片化分布只不过是全局用户中心趋同过程中的表象和局部。通过对微博用户子类的实证分析,研究发现:其一,子类越处于连接网络的中心,与全局各类用户的平均相似度越高,表现出趋同结构的中心化。其二,子类越处于相似网络的中心,与相似中心的相似度越高,表现出趋同结构的单中心性。其三,子类越处于用户子类连接的中心,也就越处于用户子类趋同的中心,并进而在类和类之间的社会连接中形成朝向共同中心的规律化趋同,表现出连接网络和相似网络的同构性。用户子类“巴尔干化”趋异和“中心化”趋同并不是冲突的而是统一的,这种中心趋向性的社交网络结构亟待进一步的理论重视。

**【关键词】**社交网络;微博用户;巴尔干化;中心化;趋同结构

**【作者简介】**徐翔,同济大学艺术与传媒学院教授、副院长;夏敏,同济大学艺术与传媒学院硕士生;唐颖,上海交通大学媒体与传播学院硕士生;王晓虹,复旦大学新闻学院助理研究员。

**【原文出处】**《新闻记者》(沪),2022.10.28~40,59

**【基金项目】**国家自然科学基金青年项目“社交网络互动中用户‘信息窄化’机理分析:基于微博的数据挖掘”(项目编号:71804126)。

## 一、问题的提出

在“同声相应,同气相求”的社交背景下,社交媒体能够使人快速找到具有相似观点的用户与群体,形成“人以类聚,物以群分”的分裂状态。当我们把社交网络用户切分为一个个群落或子类,并将视野从个体、局部放大到整个网络社会层面时,那么,社交网络究竟是碎片化的、在各个子类之间形成趋异的分化,还是有着中心化的趋同结构?这是一个重要的理论问题。

网络巴尔干化(cyber-balkanization)是学界较为主流、有着广泛影响力的观点,它指的是网络空间分

裂为不同兴趣的繁多子类,使得同一领域以及领域之间的“裂痕”越来越大(Van Alstyne & Brynjolfsson, 1996)。相似的个体之间更容易建立联系,其连接也会更加紧密;而不同子类之间差异较大,互相的连接也较为松散。与“网络巴尔干化”类似的观点还包括“文化部落主义”、“亚文化”分裂等,认为网络所带来的是双向互动和去中心的“第二媒介时代”(波斯特, 2005:3-4)。罗伯·基钦将“文化部落主义”描述为“当人们只同自己喜爱的人交流时,由此形成的兴趣社区很有可能削减多样性和窄化影响范围”(胡泳, 2016)。虽然存在着一些不同的声音,但关于子类之

间碎片化和分隔、趋异的主流观点遭受的质疑和挑战并不大,这对于我们进一步的思考有一定困难,但也留下了进一步研究的空间。“巴尔干化”的观点虽然注意到用户群体内容生产的差异性,但对碎片化背后潜在的同质化和统一性关注不足。重要的问题,不仅仅是社交网络是否处于巴尔干化的一个个孤岛之中,而是它是否处于一个单一的、巴尔干化的整体孤岛之中。在分析社交网络结构时,既不能忽视局部的碎片化、多样化,也不能对全局的趋同性视而不见,而是要将两者融入理论整体中。

因此,在社交媒体和微博时代,虽然存在着相似用户、相似观点的聚合网络所带来的区隔、封闭和割裂,但仍有几个问题需要进一步思考和明确:一是同质性驱动下的社交网络结构究竟是中心化的,亦或是去中心化的?二是在用户“越连接—越相似”的逻辑下,高连接度用户之间会产生更高的相似度,高相似度用户也更容易产生连接,结果是否并不会导致分散的“巴尔干化”,而实际上是“中心化”的趋同现象?三是“巴尔干化”和“中心化”并不是矛盾的,而是统一的,前者可能是后者的表现和表象?即碎片化的用户群体背后是否存在着趋同的统一性?四是如果存在这种统一性,其中的趋同规律和机制是怎样的?

本文提出的核心观点是:各类社交媒体用户并非巴尔干化或去中心化的,而是具有朝向中心子类聚拢和趋同的“中心化”结构;所谓的区隔化、碎片化的表象、现象,背后存在着统一性和趋同性,子类的碎片化分布只不过是全局用户中心趋同过程中的表象和局部。本文的实证检验将基于典型社交媒体之一的新浪微博样本,聚焦于用户群的社交网络“连接—相似”结构,从全局层面揭示“巴尔干化”碎片化表象下各类用户的中心趋同结构。

## 二、文献回顾和研究假设

在用户及内容同质化的研究中,“巴尔干化”在多数情况下成为约定俗成的理论背景和观点被引用。我们并不否认同质化连接网络所带来的信息孤岛,但这仅仅从局部视角强调了用户的碎片化、多样化,并不能从理论和实证上对社交网络趋同规律予以充分支撑,甚至会忽视全局结构中潜在的同质

性。因此,在研究社交网络结构时,需要打破对于局部相似人群的分析,着眼于全局,更大范围地发现用户类与类之间的关联性,将多种相关因素统合到同一个框架内,得出更具普遍意义的相关结论。尤其是对社交媒体用户及群体的连接度、相似性、中心性等因素的关联性,在现有研究中涉及较少,却是亟待关注的问题。

### (一)用户之间连接度与内容相似性的正向关联

对于社交网络结构中的用户关系,一般认为,社会关系的连接度与用户内容相似性具有正向的相互作用:即越连接越相似,同时,越相似也越连接。例如,Yu和Mo(2015)研究表明,微博用户间的兴趣相似性与他们的社交关系强度呈显著正相关。Saint-Charles和Mongeau(2018)验证了影响力网络的中心性与话语相似性网络的中心性呈正相关,即强(和相互)联系的话语相似性更高。Liang和Shen(2018)表明自我和关注对象的差异可以表现出磁铁的性质:当两个实体彼此靠近时,引力更大,并且引力与差异成负比例。可以看到,社交关系能够有效增强用户相似性,相似性也能带来更多在线交往和互动。这一方面能够导致类同于“人以类聚”或“同声相应”的碎片化的用户类;但另一方面,从社交网络全貌来看,高连接度对应的是高相似度,并且这种相似性会带来进一步的相似化,那么各类用户会被编织成一张相互连接且中心化趋同的大网。有研究也关注到了用户之间的这种相似级联,例如,Crandall等(2008)发现相似性和社会互动之间有明显的反馈效应,人们在第一次交流之前很快变得更相似,之后很长一段时间内也会变得更相似。Zhang等(2017)研究发现在线社区用户之间的同质化程度越高,不仅连接强度会随之增加,而且彼此相互影响程度也会更大,传者和受者越有可能相互接受和支持。Centola等(2007)认为同性联系的产生是因为人们与相似的人互动,随着时间的推移,他们相互影响,变得越来越相似。杨雷、吕敏(2011)研究表明子群体之间交互会促进整个群体达成共识。Lu等(2017)研究发现,在社交网络中增加子群体之间的连接数量可以加快群体意见的收敛速度,为了加速达成共识必须加强不同子群中个人的交流。同质性本质上是一种选择机制,但同

时也是在个人和群体层面的一种(潜在的)影响机制。由此可见,用户内容相似度和社会连接的紧密程度存在双向的作用关系,并且这种相似性会通过社会互动或社会影响进一步增强用户之间的趋同性,即因为相似而变得更加相似。因此从全局视角来看,在“越连接—越相似”的作用下,社交网络最终会变得更加集中化而非离散化。

## (二)网络用户观念与内容的整体趋同性

社交网络模型研究和观念动力学的提出,都表明了网络用户群体是趋同而非趋异的,群体意见和观点最终会收敛于超级影响者、公众媒体等。有研究表明随着交互的持续进行,网络社群中的个体观点将由混沌无序状态过渡到集中有序状态,不同观点的一致性程度逐步提高,最终可能形成共识,这种观点演化问题在群决策(Li et al., 2014)、少数观点的幸存(薛娟等, 2016)等社会现象中普遍存在。当传播无限继续下去时,一个封闭的系统中的所有参与者(个人、群体或者国家)都会遵从一般的集体思维模式而渐渐趋同……呈现出相似性,均衡值趋于信息所表示的集体均值(Barnett & Kincaid, 1983)。Hosanagar等(2014)指出,虽然过去的观点认为受众兴趣分散化、社群分层化会导致碎片化,但是算法推荐在一定程度上强化了小众兴趣爱好群体的娱乐需求,并且让其形成一个个的组群,导致在整体的娱乐消费市场中,用户之间变得越来越同质化;算法推荐虽然让受众形成了一个分散的小众群体,但是在整体上用户的购买特点网络越来越相似并且会逐步趋近于一个整体。高剂斌等(2018)则提出交互半径是制约全局群体行为形成的主要因素,当时间允许达到无穷时,所有个体的观念值最终都收敛到媒介观念值。

国内外学者也已提出多种模型来解释舆论的形成和演进现象,刻画了用户观点的收敛过程和影响因素。张亚楠等(2017)基于HK模型建立了社交网络舆论形成模型,研究发现在不同信任阈值下,群体观点均能收敛到一致,形成舆论共识,且信任阈值越大,收敛时间越短;同时扩大交互集合,提高人际相似性的作用强度会促进舆论共识的形成。曹春萍、康远路(2021)等利用“昆山反杀案”话题的微博用户

数据,提出了一种基于用户推荐的HK(URHK)模型,结果表明随着用户相似性阈值的增大,用户观点值达到一致的时间越快。李根强等(2017)基于改进的Deffuant模型提出,强关系更利于网络社群舆情观点的快速聚合,且观点聚合达到稳态时的观点数量和观点分布方差均小于弱关系环境中。Tavana等(2017)发现网络中具有平均偏好的标准用户有助于减少集群之间存在的分散和差异,形成更具凝聚力的网络。这些分析展现了整个社交网络内容的演化是趋于收敛而非多样化、多中心的可能性,随着相似度的提高,用户的意见交互和趋同速度也随之提高,最终呈现收敛于某一类用户的中心化结构。

## (三)核心角色的跨圈层传播与同质化扩散

社交网络中的核心角色也可能是信息跨圈层流动的突破口,如关键意见领袖、超级影响者等,他们会打破传统“巴尔干化”结构下的圈层区隔边界,并形成社交网络中连接全局、相似于全局用户类的中心结构。从传播内容和用户观点而言,网络群体的趋同化往往是指向意见领袖,即其他用户与社交网络的核心用户越来越相似;这主要是因为意见领袖代表着高影响力用户,作为连接中心拥有着更高的同化能力。例如,韩运荣、高顺杰(2012)通过社会网络分析发现意见领袖能够连接起各自离散的节点,使得原来无序、混乱的网络变得紧密。郝宇辰、刘臣(2020)发现维持高程度的领袖信任以及较多的社交联系将有助于领袖引导社区观点向领袖观点靠近。李根强、方从慧(2017)则指出意见领袖更可能成为其他个体进行观点交互的参照方,使得分散的群体观点最终朝向意见领袖所持的观点逐渐偏移。蒋巍川(2009)在考察多Agent的群体行为时提出,非平滑扩散趋同中不同Agent具有不同的趋同能力,社会地位或凸显度高的Agent聚集度高于普通Agent。从这些研究中可以发现,存在着一些关键性的用户角色或群体,他们有着与全局用户的高相似性、高连接度、高中心性,能够引导全局趋于共识,其跨圈层性对于形成中心化的网络群体结构有重要作用。由此我们认为社交网络结构在关键用户之间的强连接下,更可能具有衍生效应,展现为由高层级到低层级的中心趋向性。

#### (四)网络结构的不平衡性与集中性

“富人俱乐部”现象、“核心—边缘”结构等研究不仅探讨了用户关系,更蕴藉着社交网络结构上的不平衡性和集中性。已有研究表明,意见领袖、高连接度用户等核心用户群体不仅彼此之间有高连接度和高相似度,而且和边缘群体也具有着高连接度。例如 Zhou 和 Mon-dragon(2004)研究发现互联网在自治层的拓扑结构存在一定的富节点和“富人俱乐部”现象,即很少一部分节点具有大量的连接,而且这些节点倾向于互相连接,形成一个“富人俱乐部”。Mislove 等(2007)通过对大规模社交网站 Flickr、YouTube 等的研究发现,这些网络用户节点包含一个高度节点密集连接的核心,这个核心连接着小群强聚集的低度节点,而这些低度节点则位于网络的边缘。Borgatti 和 Everett(2000)提出了核心—边缘的结构特征,即处于核心的节点之间联系紧密,而处于边缘的节点之间联系比较稀疏,并且大多数的边缘节点与核心节点建立联系。这种结构特征也被应用于微博网络中的群体结构并得到了验证(王晓光等,2011)。Guo 等(2011)发现新浪微博用户关注关系网络中,大部分关联都是单向的,且有一个结合紧密的核心网络,其半径相对于其他社会网络来说更短。从整个传播环境来看,巴尼特和罗森(2008:120-124)曾论证了传播网络中的趋同理论,并指出“传播密度较大的局部区域共享着较多的信息,因而与密度较低的区域相比,差异或者意见的分歧以较快的速度减少”。可以看到,无论是“富人俱乐部”现象,还是“核心—边缘”结构等,都指向社交网络结构的核心节点,他们与网络中心、边缘用户类都具有高连接度,而边缘节点只和边缘的用户类具有稀疏的连接;并且连接越紧密、越相似的用户类彼此之间也更加趋同和“稠密”,而边缘用户之间则较为分散、差异较大。这与中心化的社交网络结构也是密切相关的。

综上所述,在现有的研究中,虽然各类用户的趋同现象有被提及,但是“中心化”趋同的理论架构和结构分析还未得到充分重视。对“越连接—越相似”现象的研究不能仅仅停留在表层现象和原因分析,实际上在相似用户的关系网络背后存在更深层次的趋同结构,有待更详实的提炼和分析。社交媒体群

体的同质性和趋同化是一个动态而复杂的过程,社交媒体用户子类是否具有一个中心以及碎片化背后的统一性是如何形成的,还需要将连接度、相似度等因素置于同一框架内进行探索和验证,而非局限于单一视角,这是本文主要的研究目标。

结合上述研究基础,本研究提出最为基本的假设:微博各类用户区隔化、碎片化背后存在着中心化的趋同结构。基于该基本假设,本研究将“关注—被关注”的微博用户按照内容特征划分为若干子类进行考察,通过社会网络分析方法,揭示并验证社交媒体用户子类深层次的“中心化”趋同现象与结构。

本研究以“子类”作为研究的基本单位,它指的是在某一群体中,由具有紧密联系的节点构成的小群体(约翰·斯科特,2016:134-138)。这一分析策略的必要性如下:第一,在互联网社会中,“动态网络成为一个可以被分析的基本单位,其中的所有层级(个体、组织),都可以被连贯分析”(Bennett,2003)。基于网络结构的子类、子群分析方法在意见领袖的跨圈层传播(汤景泰,陈秋怡,2020)、群体观点交互(杨雷,吕敏,2011)等多个领域都得到了应用。第二,在本研究中,验证中心化趋同规律的关键在于检验各类用户群体是否会形成单一的“整体社会孤岛”。只有对整体网络中一个个相对独立的“孤岛”,即“子类”进行分析,才能窥见用户连接关系背后的共同特征,以个体为参照则不足以说明完整的传播图景(汤景泰等,2018)。第三,个体现象是带有波动性的,群体特征则具有共同性和稳定性,因此将个体研究转变为群体研究,能够增强不同类型人群在社会学意义上的稳定性,从而在原有的随机现象之上观测到稳定性的变化和规律。

基于以上概念阐释,本文从3个关键子假设和2个推论来验证中心假设。为方便描述,下文将由用户子类两两成为关注对的连接概率构成的矩阵简称为“连接网络”,由用户子类两两相似度构成的矩阵简称为“相似网络”。其中连接网络中点度中心度最高(即与其他各类有最高的连接度)的子类简称为“连接中心”(A1);相似网络中点度中心度最高(即与其他各类有最高的相似度)的子类简称为“相似中心”(A2)。需要说明的是,连接中心(A1)和相似中心(A2)

在实际过程中存在不一致的情况,是因为两者选择的标准并不是基于完全相同的社会网络,并不影响连接网络和相似网络整体中心度及相关性的显著程度和检验结果。

由文献回顾的第1点,即“越连接—越相似”的逻辑起点可知,子类之间连接越紧密,互相之间的内容同质性就会越来越高。那么对于某一子类而言,与尽可能多的类有尽可能高的连接度,该子类也会与尽可能多的类有尽可能高的相似度,前者恰好可以用连接网络的中心度来测量。换言之,子类越处于连接网络的中心,它与全局各类就有着越高的平均相似度。由此提出H1:

H1:各子类在连接网络的中心程度,与其和全局各类的平均相似度呈正相关。

根据文献回顾的第2—4点,网络中的意见领袖具有跨圈层传播和同质化扩散的能力,所以社会网络各类用户的趋同并不是随心所欲的,而是会向高连接度、高影响力的用户类趋同。由“富人俱乐部”现象可知,高连接度的用户类之间也具有高相似度,变得越来越同质化;因此,社会网络整体会越来越收拢而非割裂,各类用户会越来越趋同于某一中心类。那么,在各子类产生趋同的相似网络中,处于最中心的相似中心会自然成为各子类趋同的某种共同“标的”。假设在相似网络中,类A不趋同于相似中心,而是趋同于另一个中心程度很高的类B,由“富人俱乐部”现象可知,类B和相似中心也是具有高相似度的,所以类A和相似中心的相似度会越来越高,各类只可能趋同于一个中心,而不是趋同于多中心。因此,各子类越处于相似网络的中心,也必然会和相似中心越来越同质化。由此提出H2:

H2:各子类在相似网络的中心程度,与其和相似中心(A2)的相似度呈正相关。

如果H1和H2均成立,在此基础上可以进一步验证连接网络和相似网络的同构性。对于某一子类来说,全局均似度越高,意味着会使越多的类与它产生尽可能高的相似程度,因而该子类成为趋同“标的”相似中心的可能性也越大,这一点恰好可以用相似网络中的中心度来测量。因此,子类越处于连接网络的中心,也就越处于相似网络的中心,由此

提出H3:

H3:各子类在连接网络的中心程度,与其在相似网络的中心程度呈正相关。

结合H2的结果,在H3假设基础上继续补充和延伸,则有子类在连接网络中越处于中心,与相似中心的相似度也越高。由此提出H3(a):

H3(a):各子类和相似中心(A2)的相似度,与该子类在连接网络的中心程度呈正相关。

根据H3,子类在相似网络中越处于中心,意味着它与连接中心就越近,与连接中心的连接度也越高,由此提出H3(b):

H3(b):各子类和连接中心(A1)的连接程度,与该子类在相似网络的中心程度呈正相关。

上述子假设和推论的关系如下,H1是H2成立的基础,H1和H2共同构成了H3的假设基础,H3(a)和H3(b)两个推论是H3的自然延伸,旨在验证相似性、社会连接度、中心度之间存在一个不断加强的社会强化过程,从而可以更好地揭示用户各类并不是区隔化、碎片化的,而是有着统一的趋中心结构。各假设之间的内在关系图如下(见图1)。

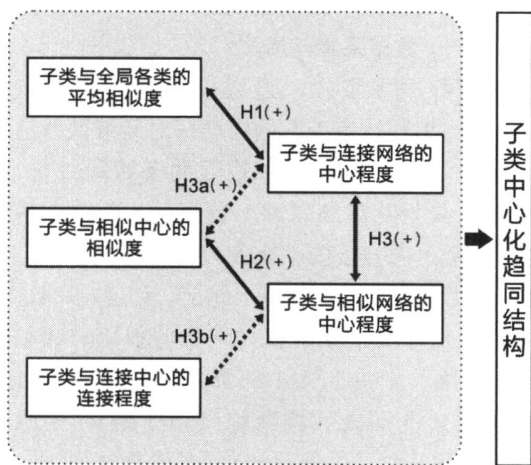


图1 子类“中心化”趋同结构及其假设的内在关系图

### 三、研究方法

本研究选取新浪微博作为研究样本平台,可参考性如下:新浪微博是我国的重要而典型的网络社交平台,微博用户之间可以通过“关注”操作建立联系,关注者成为被关注者的“粉丝”,这种关注关系是社交网络连接形成的基础。截至2021年12月,新浪微博月活跃用户数达5.73亿,同比增长约5200万;日

活跃用户数达2.49亿,同比增长约2500万(新浪财经,2022)。新浪微博用户中既有各类高影响力和高连接度的“大V”,也有形形色色的“草根”,对其中社会网络的分析具有较强的代表性和普适性。

在具体的实证过程中,本研究的实施步骤主要如下:首先,随机选取新浪微博上的数万个用户及其对应的关注者,匹配成关注对,分别抓取这些用户个人信息及所属数千万条样本帖,并进行必要的清洗和处理。其次,将这些用户所属的帖子转化为向量进行聚类,从而判定样本用户的类型和所属子类。再次,分别计算类与类两两之间的连接程度和内容相似程度,构建连接网络和相似网络。最后,利用SPSS和社会网络分析方法验证微博用户社会网络的“中心化”趋同结构。

#### (一)数据采集与清洗

在数据抓取上,通过八爪鱼采集器和python编程语言编写网络爬虫对新浪微博进行采集。采取成本相对较低的多阶段抽样,抓取工具为八爪鱼软件。首先,从新浪微博首页47个内容板块(社会、国际、科技、科普、数码、财经、股市、明星、综艺、动漫、宗教、萌宠、法律等)中,每天早晚各抓取一次帖子,持续1个月,获得65650条“种子”帖,种子帖广泛而大致均衡地分布在这47个板块。随后从种子帖做第一阶段“滚雪球”扩散,对这些种子帖抓取得到5053998条评论及其评论者,这些评论者和原种子帖发布者去重后形成3501153个用户的初始库。其后,从该初始库中随机选取并有效抓取得到87739个用户所发的帖子。之后进行第二阶段的“滚雪球”,从这87739个有效用户中,根据每人所关注的100个被关注者(新浪微博官方只开放显示5页×每页20位被关注者)随机抽2名。其后对所有用户统一时间口径和帖子数量:一是对每个用户只选取2015年1月1日至2018年12月31日期间发布的帖子;二是对每个用户,在上述这四年的时间段内一律只随机选取500条帖子,少于这个数量的用户不纳入分析。微博数据显示,2021年12月微博月活用户达5.73亿(新浪财经,2022),而早在2018年,微博活跃用户日均文字、图片、视频等发布总量就超2.5亿(新浪科技,2018)。据此计算,每个用户日均发布微博至少0.44条,年均

159条。由此看来4年达到500条的发帖总数并不算多,低于平均值,因而仍然可以保留多数有一定活跃程度的用户,并避免“僵尸用户”或长期停更用户、过度的“潜水用户”;同时,对用户需要“据其帖,察其人”,如果发帖数太少则难以了解该用户的内容特征。确保“关注—被关注”的配对双方都具有符合上述条件的帖子和有效个人资料,最终剩余13762组“关注—被关注”的“1对1”用户关系。

#### (二)用户内容的特征提取与向量化表示

在数据处理上,把单个用户的各条帖子,无顺序拼接为一个长文本,先通过向量空间模型(VSM)得到每个用户的词频矩阵。词频矩阵的获取,采取常用的模块scikit-learn中的CountVectorizer()函数,其中最低词频数设为50,最大文档频率设为0.2,只采用一元词,得到词频矩阵后转为经L2规范化的TFIDF矩阵X,转换函数为sk.learn的TfidfTransformer(norm='l2')。对于矩阵X,采取潜在语义分析(LSA)进行降维和内容特征提取(Dumais,2004)。LSA利用奇异值分解(SVD)技术把数十万以上的高维、稀疏的词频矩阵降到只有数千、数百的低维表示,并通过精简的维度提取和表达原矩阵的主要信息。

本研究中,对每个用户的词频矩阵一律通过潜在语义分析降维到1200维。降维工具选择开源模块scikit-learn中的TruncatedSVD()函数。选择1200维的维度,主要出于以下原因:TruncatedSVD()计算得到的解释方差比(explained\_variance\_ratio\_)值达0.726(见图2),用尽可能精简的维度,提取和高效保留了原

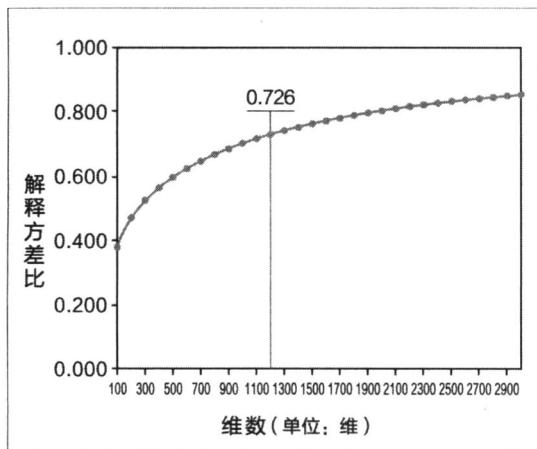


图2 潜在语义分析在不同降维维数的解释方差比

有的数万维矩阵的大部分信息,过滤掉了多数噪音或带有干扰性质的非重要信息,从而表示用户内容特征并用于用户聚类 and 内容相似度的计算。

### (三)用户内容聚类及相似性计算

基于每个用户的帖子向量,通过K-means算法进行若干种聚类类型的试验,最大迭达次数设置为6。结合误差指标(见图3),由于随着聚类数量k的增长,其簇内误差平方和(K-means Inertia)变化较为平稳,而太低的聚类数量不利于统计检验,过多的聚类数量会出现类型分布极不均衡的情况,因此根据实际的可操作性,选择将样本用户聚为200类。当k为200时,聚类效果处于平稳变化,是可以接受的选择。

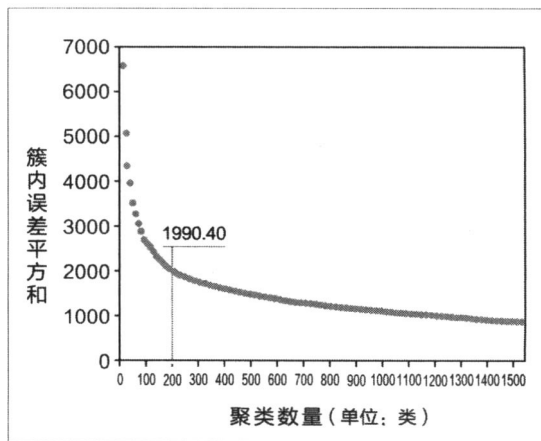


图3 聚类误差变化

用户相似度的计算采用余弦距离这一常用指标。余弦距离指两个n维点之间夹角的余弦值 $\cos(\theta)$ ,取值范围为 $[-1, 1]$ ,夹角余弦值越大,表示两向量的

相似度越高。其计算公式为:

$$\cos(\theta) = \frac{\sum_{k=1}^n x_{1k} x_{2k}}{\sqrt{\sum_{k=1}^n x_{1k}^2} \sqrt{\sum_{k=1}^n x_{2k}^2}}$$

本研究中通过类平均法计算类与类之间的用户平均余弦相似度,该值越大表明两类用户之间的相似度越高。任意两个用户 $U_m$ 和 $U_n$ 之间的余弦相似度 $\cos(\theta)$ 表示为两者之间的函数 $R(U_m, U_n)$ ,则任意一组用户 $G_1$ (包含 $n_1$ 个用户)和另一组用户 $G_2$ (包含 $n_2$ 个用户)的内容相似度表示为:

$$H(G_1, G_2) = \frac{1}{n_1 n_2} \sum_{x_i \in G_1} \sum_{x_k \in G_2} R(U_i, U_k)$$

### (四)社会网络分析:用户子类间的连接网络和相似网络的分析

社会网络分析是将行动者所处节点之间的关系用网络的形式描述出来,从而刻画他们在网络中的结构属性的方法。连接网络一般是指基于节点的连通性和连接紧密程度所构成的网络链路,能够反映整个网络的拓扑结构和密度状态,节点之间的连线代表两者存在连接关系,连线的方向代表两者的连接方向(杜洪涛等,2016;刘颖,张焕,2014)。相似网络则是基于节点内在属性与外在行为活动的相似性,并结合邻居结构和群结构等网络结构影响所构成的网络链路,两个节点的某种相似度特征越高,则这两个节点越有可能建立边的关系(马翔等,2020)。

基于“关注—被关注”的微博用户子类间的连接程度和相似程度,本研究构建了连接网络和相似网络,并选取社会网络分析法中的点度中心度和特征

表1

微博用户子类的连接矩阵(前10类示例)

|           |    | 被关注者子类(%) |        |        |        |        |        |        |        |        |        |
|-----------|----|-----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|           |    | 1         | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      | 10     |
| 原关注者子类(%) | 1  | 0.0436    | 0.0000 | 0.0109 | 0.0073 | 0.0036 | 0.0109 | 0.0000 | 0.0036 | 0.0000 | 0.0109 |
|           | 2  | 0.0000    | 0.0073 | 0.0000 | 0.0036 | 0.0000 | 0.0000 | 0.0000 | 0.0036 | 0.0000 | 0.0109 |
|           | 3  | 0.0109    | 0.0000 | 0.6976 | 0.0000 | 0.0000 | 0.0036 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
|           | 4  | 0.0073    | 0.0036 | 0.0000 | 0.0654 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0036 |
|           | 5  | 0.0036    | 0.0000 | 0.0000 | 0.0000 | 0.9156 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0036 |
|           | 6  | 0.0109    | 0.0000 | 0.0036 | 0.0000 | 0.0000 | 0.4941 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
|           | 7  | 0.0000    | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0073 | 0.0000 | 0.0000 | 0.0036 |
|           | 8  | 0.0036    | 0.0036 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0363 | 0.0000 | 0.0182 |
|           | 9  | 0.0000    | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0145 | 0.0036 |
|           | 10 | 0.0109    | 0.0109 | 0.0000 | 0.0036 | 0.0036 | 0.0000 | 0.0036 | 0.0182 | 0.0036 | 0.0509 |

向量中心性作为衡量子类中心性的指标。

### 1. 各子类之间的连接网络和相似网络

各子类连接程度以概率矩阵和交叉表的方式表示如下(见表1)。每一行表示微博原关注者的类别,每一列表示微博被关注者的类别,每个单元格值表示对应类别中形成“关注—被关注”关系的用户数量占全体用户的比例。例如,对于第1类原关注者而言,他关注的用户中有6位同样属于第一类,则1类原关注者和1类被关注者的连接比例为 $6/13762=0.0436\%$ ,以此类推,所有200类用户连接矩阵的单元格值合计为100%。由于篇幅较长,仅展示前10类用户的连接矩阵。

需要说明的是,由于连接网络具有方向性,因此“1类原关注者—2类被关注者”与“1类被关注者—2类原关注者”的连接程度不同,在后续计算中取两者均值作为1类和2类总体的连接程度。

各子类相似程度则基于余弦相似度计算结果表示如下(见表2)。每个单元格值表示所属列和所属行的子类用户帖子的相似程度。由于篇幅较长,仅展示前10类用户的相似矩阵。

以上两个矩阵即为本研究中社会网络分析的数据基础,通过Ucinet软件分别构建200个子类组成的连接网络和相似网络。

### 2. 中心度分析

本研究将采用社会网络分析的中心性指标,对用户 in 社交网络中的连接状况、相似程度和中心地位进行衡量。在社会网络分析法中,中心性测量了

行动者在社会网络中所处的核心地位和权力影响,反映行动者获取资源的优势差异,指标包括点度中心度、中介中心度、接近中心度、特征向量中心性等(约翰·斯科特,2016:91-96)。其中,点度中心度指的是在社会网络中一个行动者与其他行动者发生直接联系的能力,点度中心度越高,意味着节点在社会连接网络越居于中心位置,从而拥有较大的权力。特征向量中心性是衡量社会网络中一个节点的影响力的指标,即一个节点的重要性既取决于其邻居节点的数量(即该节点的度),也取决于其邻居节点的重要性;特征向量中心性越高,该节点连接的节点就越重要,其在社会连接网络也就越居于重要位置。总体而言,中心度最高的节点,在社会连接网络也就居于最中心和最重要的位置。

本文要衡量的标准是微博子类间直接连接的强度而非间接连接的强度,所以,本文主要考量的是“关注—被关注”的微博用户子类的点度中心度和特殊向量中心性指标,未使用中介中心度和接近中心度。基于上述连接矩阵和相似矩阵,通过Ucinet软件分别计算两个网络的点度中心度和特征向量中心性指标。

## 四、研究发现

### (一)H1的检验

根据子假设H1,对各子类在连接网络的点度中心度、特征向量中心性与其全局均似度进行Spearman相关系数及显著性的检验。结果显示均存在显著的正相关性,相关系数分别为0.729和0.697,  $p<0.01$ ,

表2 微博用户子类的相似矩阵(前10类示例)

|               |    | 被关注者子类(%) |        |        |        |        |        |        |        |        |        |
|---------------|----|-----------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
|               |    | 1         | 2      | 3      | 4      | 5      | 6      | 7      | 8      | 9      | 10     |
| 原关注者<br>子类(%) | 1  | 0.0009    | 0.0165 | 0.0036 | 0.0026 | 0.0033 | 0.0016 | 0.0062 | 0.0184 | 0.0055 | 0.0150 |
|               | 2  | 0.0165    | 0.3490 | 0.0195 | 0.0098 | 0.0152 | 0.0094 | 0.0278 | 0.2640 | 0.0216 | 0.2138 |
|               | 3  | 0.0036    | 0.0195 | 0.8328 | 0.0266 | 0.0103 | 0.0151 | 0.0301 | 0.0279 | 0.0837 | 0.0372 |
|               | 4  | 0.0026    | 0.0098 | 0.0266 | 0.1540 | 0.0246 | 0.0310 | 0.0355 | 0.0155 | 0.0122 | 0.0188 |
|               | 5  | 0.0033    | 0.0152 | 0.0103 | 0.0246 | 0.9001 | 0.0075 | 0.0219 | 0.0207 | 0.0089 | 0.0365 |
|               | 6  | 0.0016    | 0.0094 | 0.0151 | 0.0310 | 0.0075 | 0.6589 | 0.0275 | 0.0151 | 0.0091 | 0.0256 |
|               | 7  | 0.0062    | 0.0278 | 0.0301 | 0.0355 | 0.0219 | 0.0275 | 0.7831 | 0.0428 | 0.0193 | 0.0565 |
|               | 8  | 0.0184    | 0.2640 | 0.0279 | 0.0155 | 0.0207 | 0.0151 | 0.0428 | 0.5656 | 0.0337 | 0.2943 |
|               | 9  | 0.0055    | 0.0216 | 0.0837 | 0.0122 | 0.0089 | 0.0091 | 0.0193 | 0.0337 | 0.7591 | 0.0398 |
|               | 10 | 0.0150    | 0.2138 | 0.0372 | 0.0188 | 0.0365 | 0.0256 | 0.0565 | 0.2943 | 0.0398 | 0.2480 |

N=200。这表明在连接网络中,子类自身在连接网络中越趋于中心和重要位置,该子类与全局用户就越相似。也就意味着,越是高连接度的用户子类,就越是全局趋同的中心。

## (二)H2的检验

根据子假设H2,对各子类在相似网络的点度中心度、特征向量中心性与该子类和相似中心(A2)的相似度进行Spearman相关系数及显著性的检验,结果显示均存在显著的正相关性,相关系数分别为0.933和0.964,  $p < 0.01$ ,  $N=200$ 。这表明子类自身在相似网络中越趋于中心和重要位置,该子类与相似中心就越相似。换言之,各个趋同的子类只可能越来越相似。结合H1的检验结果可知,如果存在另一个中心程度很高的子类,它一定也是在全局意义上具有高相似度的子类,并且离全局相似度最高的子类越近,而不存在趋同于其他中心的可能,即各子类的趋同中心具有单中心性。

结合H1、H2的检验结果,我们可以推导出:只存在一种中心用户,它和全局各类都具有最高的相似度和连接度,并且全局各类用户都向着与全局相似度最高的中心子类趋同。同时,随着中心度的递增,子类间的相似度也由低到高地变化,越靠近中心位置,子类之间的连接和趋同性也越“稠密”。

## (三)H3、H3(a)、H3(b)的检验

通过H1、H2的实证结果,我们可以看到,这种中心趋向性和中心单极化并非偶然的、无规律的,而是基于子类之间相似度的演化规律。那么这种“中心化”趋同结构与子类之间的相似度在多大程度上有着怎样的关联和对应性,还需要进一步的检验和分析。因此,H3、H3(a)、H3(b)将“连接网络”和“相似网络”同时纳入考察,旨在探究社会网络结构更深层次的趋中心机制。

根据子假设H3,对各子类在连接网络的中心度指标与在相似网络的中心度指标进行Spearman相关系数及显著性的检验。相关性分析显示,各子类在连接网络的点度中心度与其在相似网络的点度中心度和特征向量中心性相关系数分别为0.748和0.737,  $p < 0.01$ ,  $N=200$ ;在连接网络的特征向量中心性与上述两者的相关系数分别为0.714和0.716,  $p < 0.01$ ,  $N=$

200。各子类作为连接中心的程度与其作为相似中心的程度均呈显著的正相关性。同时,在Ucinet中利用QAP相关分析进一步计算连接网络和相似网络两个矩阵之间的相关性,结果显示,两者相关系数为0.512,显著性为0.000。这表明连接网络和相似网络呈正相关,并且这种关系在统计意义上是显著的。

以上结果表明,从全局意义上而言,子类在社会连接网络和相似性网络中的中心程度具有相关性和一致性,越处于用户子类连接的中心,也就越处于用户子类趋同的中心。

根据推论H3(a),对各子类在连接网络的点度中心度、特征向量中心性与其和相似中心(A2)的相似程度进行Spearman相关系数及显著性的检验。结果显示均存在显著的正相关性,相关系数分别为0.783和0.807,  $p < 0.01$ ,  $N=200$ 。根据推论H3(b),对各子类在相似网络的点度中心度、特征向量中心性与其和连接中心(A1)的连接程度进行Spearman相关系数及显著性的检验。结果同样呈显著的正相关性,相关系数分别为0.617和0.630,  $p < 0.01$ ,  $N=200$ 。这表明中心度、连接度、相似度存在互相强化的关系,再次验证了微博用户连接网络与相似网络的同构性。综合H3、H3(a)、H3(b)的实证结果,无论是连接网络还是相似网络,微博用户各类并不是离散化的,而是存在着趋同的单中心,并呈现在类和类之间的社会连接中形成朝向共同中心的规律化趋同。

## 五、结论和讨论

基于“关注—被关注”的13762对微博用户样本,本文考察并检验了微博用户“中心化”的趋同现象与结构。研究发现社交媒体用户子类的网络结构并非是碎片化、分散化的,而是呈现“中心化”的趋同结构和趋中心结构。具体而言:(1)趋同结构的中心化:用户子类越处于连接网络的中心,与全局各类用户的平均相似度越高,两者之间存在显著的正相关性。(2)趋同结构的单中心性:用户子类越处于相似网络的中心,与相似中心的相似度就越高,两者之间存在显著的正相关性。(3)社会连接和相似网络的同构性:用户子类越处于连接的中心,也就越处于趋同的中心,两者之间存在显著的正相关性。具体表现为:与相似中心越相似的子类,就越处于连接网络的中

心;与连接中心连接越紧密的子类,也越处于相似网络的中心。

本文对微博样本的实证分析,推进了对社交网络“巴尔干化”和“中心化”的现有认识,对传播学理论发展具有一定启发意义。

其一,打破了社交网络“巴尔干化”的惯性认知。当前对巴尔干化、去中心化等主流观点的研究中,或是过度强调社交网络碎片化、区隔化的分布特征(Vicario et al., 2017),或是仅强调局部、小范围用户个体的同质性(Kim, 2019),对全局相似度和宏观的趋同结构重视不足。从更广泛的传播生态来看,无论是社会连接还是社会影响,所带来的同质化不仅是相似用户的局部连接和聚集,更是全局意义上的相似结构演化,而去中心化、碎片化只是形式上的表现(董璐, 2012)。本研究对不同用户类型之间的全局趋同性进行了更深层次的剖析,指出社交媒体各类用户之间并不会形成碎片化、区隔化的“部落”,而是存在“中心化”的趋同结构和趋中心结构,推进了对于“网络巴尔干化”的认识。

其二,探究了“巴尔干化”和“中心化”的统一性。一方面,部分学者认为在多元化的社交信息中,异质性是社交网络用户的主要特点(Brundidge, 2010),由此形成的巴尔干化网络更加验证了第二媒介时代(波斯特, 2005: 3-4)、文化部落主义(胡泳, 2016)等理论观点。另一方面,对立的观点认为互联网便利了选择性接触和近似观点的强化,同质化的增加呈现的是高度中心化的社交网络(Freelon et al., 2015; Garrett, 2009)。虽然本研究证明了社交网络各类用户之间并不会产生趋异的分化,但这并不意味着“巴尔干化”和“中心化”是非此即彼的关系,这是已有研究容易忽视却亟待重视的理论维度。实际上,“巴尔干化”趋异确实存在,但只是全局用户类“中心化”趋同中的某一环节或表层现象,并非最终结果;两者并不是矛盾对立的,而是同一事物的不同表现,具有统一性。本研究为“巴尔干化”分化表象背后的同化趋向提供了实证支持,明确了“巴尔干化”和“中心化”是社会网络结构的碎片化表象与中心趋同本质的关系,是可以整合到一个理论整体中的。

其三,提出了不同类型用户在社会连接中形成中心化趋同的规律。在用户群体趋同过程中,意见领袖的跨圈层传播(汤景泰, 陈秋怡, 2020; 詹骞, 2014)、网络结构的不平衡性(Zhou & Mondragon, 2004; Halberstam & Knight, 2016)等往往是关注较多的影响因素。本文区别于现有研究之处在于,从群体社会连接度和内容相似性的角度分析了趋同结构的关键线索,探究了网络社会整合的途径和可能。也可以说,“巴尔干化”和“中心化”的统一性也同样隐含在社会连接的维度中。在连接度、相似性和中心度的不断强化作用下,社交网络会逐渐收敛于共同的趋同中心并趋于稳定,完成从“去中心”、“多中心”向“单中心”的过渡和统一,形成更加整合而非碎片化的趋同网络空间。

社交网络“中心化”背后的趋中心作用机制也让我们重新审视网络社会的用户实践。从用户主体角度而言,“中心化”趋同结构指向了“单向度”、“标准化”主体再生产的可能。随着用户社交媒介使用程度的提升,群体之间的差异性和独特性会不断被消磨,表现出不同程度的趋同与重复。这暗含了网络社会的未知风险,人们的思想内容可能会被禁锢在有限的空间内,形成全局性的“社会回音室”。从网络文化角度而言,网络亚文化的生产并不是随心所欲地形成碎片化的若干局部,在“中心化”趋同结构下,若干亚文化通过同质化传播扩散构成了网络文化整体逻辑的生成张力。因此,重要的问题不仅是网络空间存在怎样的亚文化,而是众多亚文化是否具有深层分化、整合的土壤,这可以成为网络亚文化研究未来的方向和角度。从网络意见和观念角度而言,不同类型用户连接度和内容相似性的社会强化呈现了网络社会思想观念趋于统一、整合的可能性,启示我们社会连接结构对网络意见结构、网络文化结构具有改变和塑造作用。这既是可以进一步剖析的带有规律性的现象,又是需要我们重视的媒介化社会问题。

本文在探讨社交网络趋同结构方面做了一定的尝试,但也存在一定的局限性。实证分析和结论尚未将动态演化过程纳入考量,全面展现中心化趋势随时间演化而加强的特征,有待后续研究基于更大

规模样本,并引入时间序列方法加以印证和补充。

#### 参考文献:

[1]曹春萍.康远路(2021).融合用户相似性的 Heggelmann-Krause 模型研究。《小型微型计算机系统》,(11),2344-2349。

[2]董璐(2012).从全球社会安全视角解析微博国际传播的高度趋同性。《国际关系学院学报》,(3),101-107。

[3]杜洪涛,孟庆国,王君泽(2016).基于社会网络分析的微博社区网络结构及传播特性研究。《情报学报》,(8),838-847。

[4]高剂斌,贾贞,郝佳佳(2018).基于个体观念演化驱动的社交网络群体行为建模及分析。《桂林理工大学学报》,(1),160-167。

[5]韩运荣,高顺杰(2012).微博舆论中的意见领袖素描——一种社会网络分析的视角。《新闻与传播研究》,(3),61-69,111-112。

[6]郝宇辰,刘臣(2020).媒体干预及用户信任的网络社区观点演化研究。《计算机工程与应用》,(16),156-164。

[7]胡泳(2016)。《文化部落主义的迁移性》。检索于 <http://www.infzm.com/content/121341>。

[8]蒋巍川(2009).多 Agent 群体行为中的扩散趋同综述。《模式识别与人工智能》,(6),877-883。

[9]李根强,方从慧(2017).复杂网络视角下网络集群行为主体的观点演化研究。《情报科学》,(5),125-131。

[10]李根强,罗艳艳,臧学莲(2017).基于有界信任模型的网络社群舆情观点演化研究。《情报科学》,(6),63-68,144。

[11]刘颖,张焕(2014).基于社会网络理论的微信用户关系实证分析。《情报资料工作》,(4),56-61。

[12]马翔,徐名海,吴宇鹏(2020).基于相似性的社交网络结构建模。《计算机应用与软件》,(12),83-89。

[13][美]马克·波斯特(1995/2005)。《第二媒介时代》(范静哗译)。南京:南京大学出版社。

[14][美]乔治·巴尼特,德文·罗森(2007/2008)。互联网在全球范围的影响:挑战与前景。《全球传播(第二版)》(叶海亚·R.伽摩利珀编,尹宏毅译)。北京:清华大学出版社。

[15]汤景泰,陈秋怡(2020).意见领袖的跨圈层传播与“回音室效应”——基于深度学习文本分类及社会网络分析的方法。《现代传播(中国传媒大学学报)》,(5),25-33。

[16]汤景泰,星辰,高敬文(2018)。论“一带一路”国际话语权提升——基于首届“一带一路”国际合作高峰论坛 Twitter 传播数据的分析。《新闻大学》,(5),56-65,148-149。

[17]王晓光,袁毅,滕思琦(2011)。微博社区交流网络结构的实证分析。《情报杂志》,(2),199-202,207。

[18]新浪财经(2022)。《微博发布 2021 年第四季度及全年财报》。检索于 <https://finance.sina.com.cn/stock/usstock/c/2022-03-03/docimewiws3985845.shtml>。

[19]新浪科技(2018)。《微博赋能内容作者收入 268 亿内容生态更加活跃》。检索于 <https://tech.sina.com.cn/i/2018-12-21/docihmutuee1410472.shtml>。

[20]薛娟,丁长青,陈莉莎,张志武(2016).基于 SIR 的众包社区知识传播模型研究。《科技进步与对策》,(4),131-137。

[21]杨雷,吕敏(2011)。子群体间交互对动态群体决策观点收敛时间的影响研究。《工业工程》,(4),104-109。

[22][美]约翰·斯科特(2012/2016)。《社会网络分析法(第3版)》(刘军译)。重庆:重庆大学出版社。

[23]詹骞(2014).政务微博意见领袖的社会网络分析——以北京地区政务微博为例。《现代传播(中国传媒大学学报)》,(12),117-122。

[24]张亚楠,孙士保,张京山,尹立航,闫晓龙(2017).基于节点亲密度和影响力的社交网络舆论形成模型。《计算机应用》,(4),1083-1087。

[25]Barnett, G., & Kincaid, D.(1983). Cultural convergence: A mathematical theory. International and Intercultural Communication Annual, 7, 171.

[26]Borgatti, P., & Everett, G.(2000). Models of core-periphery structures. Social Networks, 21(4), 375-395.

[27]Bennett, W. L.(2003). Communicating global activism: Strengths and vulnerabilities of networked politics. Information Communication & Society. 6(2), 143-168.

[28]Brundidge, J.(2010). Encountering ‘difference’ in the contemporary public sphere: The contribution of the internet to the heterogeneity of political discussion networks. Journal of Communication, 60(4), 680-700.

[29]Centola, D., Gonzalez-Avella, J., Eguíluz, V., & San Miguel, M.(2007). Homophily, cultural drift, and the co-evolution of cultural groups. The Journal of Conflict Resolution, 51(6), 905-929.

[30]Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., & Suri, S.(2008). Feedback effects between similarity and social influence in online communities. Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, 2008, 160–168.

[31]Dumais, S. T.(2004). Latent semantic analysis. Annual Review of Information Science and Technology, 1, 188–230.

[32]Freelon, D., Lynch, M., & Aday, S.(2015). Online fragmentation in wartime: A longitudinal analysis of tweets about Syria, 2011–2013. The Annals of the American Academy of Political and Social Science, 659(1), 166–179.

[33]Garrett, R. K.(2009). Echo chambers online? Politically motivated selective exposure among internet news users. Journal of Computer-Mediated Communication, 14(2), 265–285.

[34]Guo, Z., Li, Z., & Tu, H.(2011). Sina microblog: An information-driven online social network. 2011 International Conference on Cyberworlds. Washington, 2011, 160–167.

[35]Halberstam, Y., & Knight, B.(2016). Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter. Journal of Public Economics, 143, 73–88.

[36]Hosanagar, K., Fleder, D., Lee, D., & Buja, A.(2014). Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation. Management Science, 60(4), 805–823.

[37]Kim, S.(2019). Directionality of information flow and echoes without chambers. PLoS ONE, 14(5), 1–22.

[38]Li, W., Tang, S., Pei, S., Yan, S., Jiang, S., Teng, X., & Zheng, Z.(2014). The rumor diffusion process with emerging independent spreaders in complex networks. Physica A, 397, 121–128.

[39]Liang, H., & Shen, F.(2018). Birds of a schedule flock together: Social networks, peer influence, and digital activity

cycles. Computers in Human Behavior, 82(5), 167–176.

[40]Lu, A., Sun, C., & Liu, Y.(2017). The impact of community structure on the convergence time of opinion dynamics. Discrete Dynamics in Nature and Society, 2017, 1–7.

[41]Mislove, A., Marcon, M., Gummadi, K., Druschel, P., & Bhattacharjee, B.(2007). Measurement and analysis of online social networks. Internet Measurement Conference: Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement. San Diego, 2007, 29–42.

[42]Saint-Charles, J., & Mongeau, P.(2018). Social influence and discourse similarity networks in workgroups. Social Networks, 52, 228–237.

[43]Tavana, M., Santos-Arteaga, F., & Di Caprio, D.(2017). The effect of preference similarity on the formation of clusters and the connectivity of social networks. Computers in Human Behavior, 72, 208–221.

[44]Van Alstyne, M., & Brynjolfsson, E.(1996). Could the internet balkanize science. Science, 274(5292), 1479–1480.

[45]Vicario, M. D., Zollo, F., Caldarelli, G., Scala, A., & Quattrociocchi, W.(2017). Mapping social dynamics on Facebook: The Brexit debate. Social Networks, 50, 6–16.

[46]Yu, Y., & Mo, L.(2015). Investigating correlation between strength of social relationship and interest similarity. International Conference on Computational Social Networks. Beijing, 2015, 172–181.

[47]Zhang, C., Lu, T., Chen, S. & Zhang, C.(2017). Integrating ego, homophily, and structural factors to measure user influence in online community. IEEE Transactions on Professional Communication, 60(3), 292–305.

[48]Zhou, S., & Mondragon, R. J.(2004). The rich-club phenomenon in the Internet topology. IEEE Communication Letters, 8 (3), 180–182.