

属性与关系的再认识

——社会网络分析研究现状与演进

庞云黠

摘要:社会网络分析方法(SNA)的优势在于对“关系”变量的采纳与分析,改变了以“属性”变量为研究核心时抽离研究对象背景的缺点,但较少对“属性”变量进行分析,也可能成为社会网络分析方法的一个弱点。随着组织传播学者提出重视结合“关系”变量与“属性”变量进行多层次分析,经考察近些年SNA在传播学各研究领域的应用,分析其对于“属性数据”与“关系数据”研究的不同角度。研究发现:早期的SNA研究侧重个体网,属性数据与关系数据通常彼此分离;组织传播学者引入了ERGM和SIENA模型,开始尝试将两类数据置于同一模型内分析,并开始同时关注关系效果与属性效果;全网环境下的整体网研究,则体现了数据先行的优势,通过对数据的量化分析,发掘点属性在结构形成中体现的作用,少数研究也会采用ERGM等模型对两类数据进行共同分析。总体来看,SNA近些年开始较为强调关系变量与属性变量的共同探索,研究方法虽然支持两类数据逐渐结合,但是,这种结合的效果尚不理想,未来需要更多共同解释两类变量的研究工具与研究理论的出现。

关键词:社会网络分析;研究进展;属性变量;关系变量

中图分类号:G206 **文献标识码:**A **文章编号:**2096-5443(2019)03-0117-12

项目基金:国家社会科学基金青年项目(14CXW025)

现阶段国内传播学研究已经开始大量使用社会网络分析法(social network analysis, SNA),可以看到学者们面对大数据的挑战,对于新的、具有解释力的研究方法的期待。

社会网络分析方法之所以在近些年被大家广泛使用,最核心的吸引力在于它的分析对象从以前“绞肉机”似的“属性数据”(attribute data)^[1]转到了具有一定结构要素的“关系数据”(relational data),在社交网站广泛使用的今天,SNA无疑非常具有吸引力。

但是如果认为SNA的分析仅仅是对于关系数据的分析就比较片面了,事实上,国外传播学者对SNA的使用经历了一个渐进的演化过程。早在2011年Miller等^[2]在对组织传播的量化研究方法发展历史进行综述时就提到:在组织传播研究中,从20世纪70—80年代,社会网分析法主要被用在研究个体在社会网中所处的位置如何影响人们对他的认知、属性、行为的理解,因为这期间,如果关注社会网络结构的话,主要能够讨论三个方面的指标:结构对等性(structural equivalence)、直接关系(direct relations)以及中介性(brokerage)。到了20世纪90年代,研究者已经不满足于简单的研究关系数据,而开始同时对关系数据和属性数据进行考量,看他们如何共同影响研究对象。当然这种转变也是因为有了新的分析工具,主要包括:指数随机图模型(ERGM, exponential random graph models),主要解决同时考量关系变量和属性变量的问题,实证网络分析仿真研究(SIENA, simulation investigation for empirical network analysis),主要侧重研究个体驱动的动态网络演化,也可以同时考量关系变量和属性变量。

沿着上述学者关于“关系数据”与“属性数据”的结合的期待,笔者希望考察的是,如果单纯从组织传播的视角来看呈现了上述的方法变迁,两类数据走向结合,那么扩展到传播学其他类型研究,SNA的使用现状是一种怎样的状态?这种状态从方法上或是理念上体现了什么新趋势?

一、个体网研究:属性与关系的分离

超越组织传播的研究范围来看,自上述综述的2011年以来,“属性”与“关系”分离的研究仍是个体网研究当中的主流。

社会网络分析主要有两大类分析对象,一类是整体网(whole network),一类是个体网(ego-network)。整体网主要研究网络中的结构要素,主要的研究指标包括:网络密度、中心度、成分、派系等等,个体网主要分析以特定的行动者为研究中心,只考虑与该行动者相关的联系,主要研究中介性、个体网络同质化、多元化、关系强度等等。

在大量的个体网研究当中,关系数据既可以做研究的因变量,研究个人的属性如何影响这种“关系”变量;也可以作为自变量,同时结合个人的“属性”数据,共同解释其他的研究变量。注意这里虽然同时考察属性和关系变量,但他们是彼此独立的,即不同于开篇提到的将两者结合进行研究的情况,这个内容在第二部分探讨 ERGM 的应用前提时一并论述。

个体网较为常见的“关系”变量包括:关系强度、彼此之间的互惠性、亲密性、同质化或多元化等,也包括关键的指标比如中介性以及闭合程度等(brokerage and closure)。

具体来看,以“关系”为因变量,“属性”为自变量的研究,可以分析个人的性格(内向或外向等)如何影响个人讨论网的多元化程度以及个人政治参与的积极性问题^[3],或者网络使用是否会改变个人讨论网络的多元性问题,因为线下讨论网络的关系一般是同质性的,那么网络使用增加了多元化接触的机会是否会改变多元化个人讨论网络的性质^[4]?当然这里的“关系”变量可以不仅存在于个人中,也可以存在于国家间,探讨国家间在脸书平台上的友谊关系构建被哪些因素所决定,最后发现有效的影响因素包括:是否具有共同的边境,语言,文明类型以及移民流动^[5]。

另外,在社会网的研究中,很早即关注了因种族隔离带来的空间上社会隔离的问题,因为同样种族的人通常更喜欢聚居在一起,但是新技术产生之后能否突破这种空间上的限制,形成多元化的跨种族之间的沟通,就成了一个非常有价值的问题。例如,有学者研究了手机的使用对种族隔离的效果,最后发现,是形成多元化还是保持社会分层,最终要看个体的属性,主流群体通常通过 ICT 使用维持和加强自己的组内沟通(in-group),强化已有的优势,证明了社会分层理论;而少数派群体则是通过 ICT 的使用加强自己的多元化沟通(out-group),证明了社会多元化理论^[6]。

更多的研究集中在以“关系”为自变量的研究当中。

如以“关系”强度为自变量的研究,可以分析诸如强弱连带对使用者的志愿行为、慈善捐助行为态度的影响程度^[7],对个人的社会耐受度(social tolerance)的影响^[8]。除了强度以外,个体网还会单独讨论个体网中某个点的中介性和闭合程度,把他们作为重要的自变量,Shen 等检验个人游戏网络的中介性和闭合程度在进行 MMOG(多人线上游戏)过程中,如何影响他们游戏任务的完成水平以及和其他玩家的信赖程度^[9]。

在政治传播的研究中,个体网研究的核心集中在个人讨论网络(discussion network)的研究中,主要关注群体讨论的同质性或者多元化的问题。例如,讨论多元化和同质化是否会影响到公民政治参与(political participation)程度的问题^[10],讨论网络多元化和意见极化程度关系的研究也是这个领域中主要关注的内容之一^[11]。

可见,在个体网研究中,无论是属性变量、关系变量分别成为自变量、因变量,还是两者共同作为独立的自变量,两者都是首先被认为是相互独立的两个要素,这种独立虽然会被后面的研究提出一定的质疑,但是不可否认的是在 SNA 普及之后,它依然有其自身的意义,而且这类研究依然活跃,毕

竟个人网的存在状况是整体网的基础。

二、组织传播侧重历时性研究

在总结和展望未来的组织传播研究趋势与重点时, Monge 等提到了两个大的方面:一是组织的历时性发展与转变(change and development over time);二是分属不同层次的个体、小的组织单元、组织之间的彼此关系研究(relationships among people, units, organizations, and meanings),即他一直强调进行多层次(multilevel)结合的研究^[12]。

Monge 等人提到的应该把研究的侧重转移到历时性研究上,已经在近七年的研究中得到了充分的体现^[13]。可以说近七年的组织传播研究很多使用的都是历时性数据,进行的都是网络变迁的研究。

不过在研究历时性变迁的过程中,属性变量和关系变量常常被结合起来考察,这一点与上述个体网络相对独立分析两者形成了鲜明的对比。能够把两类变量结合起来的重要原因是新的分析工具的出现:ERGM 和 SIENA。这两个模型除了能够解决点属性和关系属性结合分析以外,还有一个重要的特点在于可以处理彼此非独立的个体,因为社会网络数据本身是有关联的,而以前建立在个体独立性假设基础之上的“标准”统计方法(例如回归分析等)用在这里是不合适的^[14]。

这里需要强调的是 ERGM 和 SIENA 在对关系变量和属性变量进行分析时,会把两者分别转为两类变量:一类是内生变量,一类是外生变量。内生变量实际上就是网络中的关系变量,例如网络中心度、互惠性、传递性、中心势和网络密度等等,它们本身被用来阐释关系的结构趋势;而外生变量则是网络个体的属性,是除了关系属性以外的其他属性,包括成员属性(如年龄、性别等),这里的内生变量不可由外生变量预测,而两者都用来解释网络的结构趋势^[14]。无论是内生还是外生变量,都是多层次(multilevel)的:内生变量可以包含个体层面的中心度,二元层面的相互性、互惠性,三元层面的传递性、循环性;而外生变量可以包含个体层面的年龄、性别,二元层面的属性相同性,比如同年龄、同党派等等。

两者影响网络结构的方式可以通过具体的计算来实现:即对比某个变量影响下的网络中建立关系的概率是否高于没有该变量时。

以内生变量的点中心度为例,实际网络当中具有更高中心度的点,更有可能与其他点建立连接,而这个计算只需要通过对比该点的实际形成连接比率与随机产生的网络中的点比率进行对比即可,如果对比结果显著,则能够说明中心度这个内生变量能够影响网络构建。如图 1,随机网络的获取方式是:设定与实际观测网络相同的网络大小、密度、度数分布值,然后获取 1000 个随机网络,图中的点为观测值(即实际值),箱体为随机网络的相应数值范围,如果点的取值在箱体图外,则表明该值具有显著性。

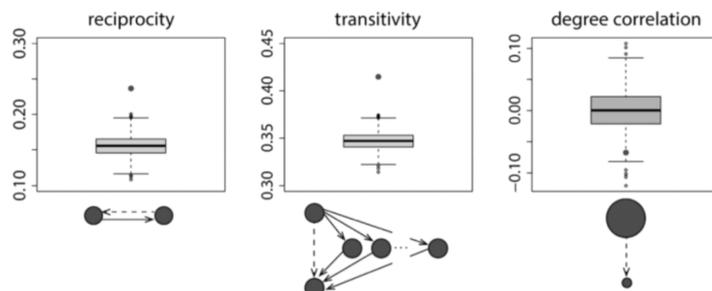


图 1 观测网络与随机网络之间的差异^[15]

具体来看较为典型的组织传播相关研究,如有学者探讨与 HIV 有关的 NGO 网络变迁过程,研究他们在选择对外链接的过程中,是如何保留或者放弃相应的链接,不同的 NGO 属性如何影响这种链接的选择与变化^[16]。

Margolin 等则着眼于组织之间的一些重要事件的发生,比如新的规范的确立,对组织之间的关系构建产生怎样的影响。例如新颁布的规范性内容对 NGO 网络结构的影响,看组织会如何调整他们和老牌机构或者新兴机构之间的关系,会更注意保留或者切断哪些类型机构的联系,网络结构的互惠性或者倾向性(preferential attachment)是否会改变等,使用的分析模型是 SEINA。^[17]

更为明确的对于两类变量的探讨体现在下文的分析当中。

三、控制关系变量之后的属性、关系效果对比

关系变量与属性变量进行结合分析的典型例子反而并不是体现在两者的共同效果上,而是体现在学者开始逐渐区分出“关系变量”与“属性变量”两者不同的效果中,下面我们来看两个递进的研究。

这两个研究的核心初衷都是一致的,即认为此前新闻传播学院的排名一般是依赖教师的研究论文数量等指标,但他们提出一个新的评估新思路:即用博士生毕业的流动情况来衡量这些院系的相关博士生项目(doctoral program)的质量好坏,主要考察的是各个项目输出(student placement)和引入博士(faculty hiring)的情况,综合考虑两者来最终评定一个项目的好坏。

第一个研究由 Barnett 等完成^[18],研究发现,对外输出学生时,各种中心度,诸如 closeness centrality, in-degree centrality, eigenvector centrality 高的,学校的信誉度(reputation)也的确比较高,似乎用输出能力评价学校的信誉度是可行的。

但 2016 年,宾夕法尼亚大学的几位学者把 Barnett 的研究深入下去,最后却呈现了完全不同于前述研究的结论。他们研究对象仍然是传播学领域的 PHD 项目在雇佣(hiring)和输出(placement)过程中形成的网络,他们的最初目标也是希望找到一套不一样的标准,能够更客观、更有效地评价现有的博士生项目。^[15]

文章分析了 2007—2014 年 7 年之间的博士毕业生的流动状况,但是最后发现教师雇佣(faculty-hiring)并非简单地反应博士教育的质量,这种雇佣的人才流动,其实具有内在的惯性。也就是说一些网络的“内生变量”,包括网络的互惠性、连通性(transitivity)以及累积优势(cumulative advantage)是影响网络结构的主要因素;而一旦控制了这些内生变量,则发现其他的“外生变量”,包括传统的教学质量评价体系中重要的,如机构的美誉度(institutional prestige)、师资的年资(faculty seniority)、学校的排名情况等虽然部分要素仍会影响网络构建,但效果相比来说十分微弱,只能被认为是影响博士雇佣与流动的次要因素,文章使用了 ERGM 模型进行上述分析。

所以文章结论指出,他们挑战(challenge)了早先 Barnett 等的研究结论,博士的输入与输出并不是教学质量的反映,而主要源于结构的内生动力,人们在雇佣师资的过程中可能更多的是依照此前的经验,身边的朋友推荐,或者自己的个人网络,雇佣行为更多的是一种内部信息的流转。

这两个连续的研究恰恰说明了本文在第二部分谈到的问题,内生变量不可以由外生变量来预测,在个体彼此具有较强相关性的前提下,传统的统计工具可能会得出错误的结论,这也是 ERGM 等模型应用的价值所在。

这种逐渐把内生变量和外生变量进行区分的研究,在近三四年中开始逐渐丰富起来。比如有研究发现相比以前提出的个体的性格、动机、能力、社会政治地位等因素,一般的讨论以及网络相关的内生性因素(general discussion and network-endogenous factors)是个人讨论网构建以及互动模式的更好的预测指标,文章也使用了 ERGM 模型进行了上述关系属性与点属性的分离,分别讨论了内生变

量与外生变量的作用。^[19]

还有研究把网络结构的影响效果从个人是否使用社交媒体中分离出来,他们希望能够解释个人在网络中获得社会支持(social support)的感知,是因为个体在不断使用社交网络(比如在脸上更新自己的状态或者与他人私聊),还是因为个人的网络结构本身?结果发现,如果控制了网络结构这个内生变量,社交媒体的使用行为中脸书状态更新(status updates)仍然与社会支持(social support)的感知呈现正相关,但是行为使用中的私聊行为(private messaging)原本与情感的社会支持感知是显著正相关的,但是控制了网络结构之后,这种显著性消失了。所以作者提出应该重视网络结构本身对于社会支持感知的影响。^[20]

较为典型的组织传播研究也开始类似的探讨。如关注组织内众包行为,针对雇员网络以及创意网络两类网络,从结构上分别探讨网络结构效果(结构上的马太效应)与点属性(包括地理同质性、是否同为高级职员等)对于创意网络讨论情况的影响。研究发现,从结构上来讲,存在显著的马太效果:一些创新度超级活跃的雇员的中心度很高,中心度很高的创意吸引了大多数雇员的注意。而从点属性来看,如果控制了上述的关系属性,则地理同质性与支持相同创意的相关度虽然显著但是微小;而资历高的人并未比资历浅的人的更愿意在众包平台上表达支持某个创意,也就是说该相关缺乏显著性^[21]。

上述各个领域对于关系变量与属性变量进行分离分析的研究似乎在表明,结构本身具有自己的内在动力与逻辑,虽然个体在构建网络结构时依据了自身属性,但是在这种网络构建完成之后,关系属性可能就具有了更强的反作用,甚至超越个体属性本身。

当然上述研究也表明,并不是每一次对内生的变量控制之后,所有的点属性都会变得不显著,也说明这种结构的内生性需要进一步深入探讨。

四、全网环境下的整体网相关研究

之所以本部分确定为全网环境下的整体网研究,与前几个部分的区分主要体现在两方面:一方面是数据来源,前述研究的数据多采用个体报告的方法,部分组织传播研究采用二手数据,而本部分的数据则多来自网络直接抓取;另一方面是研究视角,本部分主要研究侧重整体网的结构研究。

当然数据和视角都是相对而言的,因为有些研究的数据来源方式其实很多元,网络抓取、问卷调查、深度访谈等等都会出现,但网络抓取是主要的数据来源方式;研究视角也可能既有个体网也有整体网两方面的探讨,但是整体网视角为主要考量。

(一)结构化网络聚类与属性挖掘

全网环境下的整体网分析中,较多的研究集中于最终形成的网络结构,尤其是网络聚类(cluster),这种结构本身是整体网分析最关注的问题,使用的方法通常是社会网络分析软件中的各种聚类分析,当然这种分析的工具正在不断演进中。

较早的该类型研究多是对超链接进行分析,如有两篇全面描述博客网络的文章,一篇是描述阿拉伯世界的博客^[22],另一篇则是分析古巴的博客^[23]。在对阿拉伯世界的博客分析中,文章通过分析发现博客形成小团体(clusters)的界限标准,首先为国界,大体分为埃及、黎巴嫩、叙利亚、沙特阿拉伯以及科威特,此外还有一个圈落为桥群体,主要是既说阿语,又说英语的群体,因此语言成为第二个边界形成要素。而在每个国家的内部圈落形成的情况又各不相同。其中,埃及为最大的圈落,主要分为世俗改革派(secular reformist sub-cluster)、广泛的反对派(wider opposition sub-cluster)、埃及青年组织(Egyptian Youth)等等。

更多的研究则集中在政治党派个体之间的关系状态,看两者是否会形成同质化的结构圈子。典型的研究如有韩国学者研究了韩国政治领导人的线上可见(visibility)网络,选择了18th National As-

sembly 的 278 名成员,分析了他们之间网页上的共现(co-occurrence)网络,共现网络出现了明显的结构极化。^[24]

群体的聚类研究也可以体现在某个具体的事件中,学者研究 2008 年韩国网上出现的反对进口美国牛肉的抗议事件,最后发现,虽然开始的时候不可避免地出现两个对立的群体,但是后来随着主要的关键性 blog 开始提供中立的意见,导致了最后中立群体的不断扩大,因此文章认为网络公共空间是可控的^[25]。

可以看到,SNA 的整体网结构研究中,尤其是超链接的研究中,“属性”变量也可以被看作类似为因变量,或者类似数据挖掘后呈现的一种研究发现,是基于分析工具进行聚类之后产生的研究结果,而不是对事先假设的验证,比如研究最终发现了政党、文化、国家、正反两种意见的不同圈落。

需要注意的是,上述研究都是发现了不同群体形成各自的聚类的。同样是研究政治派系,或者话题冲突,也会出现完全相反的结果。

例如,Dvir-Gvirsman 的研究对象是以色列的普通民众,在搜集了民众选举前七周的网站访问数据(web logs 直接获取的行为数据)后,他发现以色列民众对于不同意识形态的网站关注(左翼或者右翼)基本不存在差异,人们的网站浏览是非常多元化的,并没有出现回音室效果,即没有选择性注意的情况;通过社会网分析方法获得了访问各类网站的社会网络图,也并没有出现类似美国的分立格局。^[26]

与韩国发现的事件讨论中中立群体的扩大不同,一个形成鲜明对比的研究关注的是政治传播中的谣言传播问题,最后发现 Twitter 没有很明显的自我净化的功能,因为在同质化的网络结构中,传播者会形成一个党派化(partisan)的结构,核心的群体只传播对手的负面信息,而且拒绝更正谣言;相反,拒绝谣言者没有办法形成一个明确的团体结构,或者形成一个集中的社群,所以从整体来看,谣言的拒绝率只有 3.37%。^[27]

产生上述的研究结果差异,是因为两对研究分别选择分析了不同的群体属性(普通民众 VS 政治精英),不同的话题属性(政治谣言 VS 事实选择),因而造成最后的结构分析出现了不同的结果。这些对于群体属性或者话题属性的控制与选择其实并不出现在研究本身当中,主要是学者们在选题时的考量因素,因此只是在这里作为一个与本文相关的内容稍做讨论,这种属性体现的更多的是一个研究语境,或者研究背景问题,但是无疑是非常重要也容易被忽略的一种“属性”数据。

回到结构化的呈现与属性挖掘,类似的研究还有从受众中心角度,探讨受众中心网络与互联网基础设施结构(technical infrastructure)网络的不同,最后研究发现后者的网络是中心化的,而前者的网络是去中心化的,受众中心网络主要依靠语言进行区分,文化的区隔是核心^[28]。同样从受众中心的角度出发,有研究对新媒体时代受众碎片化结构提出了驳斥,作者以观看的节目为节点(node),超过一定比例的重复受众则会构建一个连接,最后通过社会网络分析发现,受众在观看 236 个媒体节目过程中,有很大一部分是重复的,而并没有形成所谓的孤立的忠实受众。因此作者对不少文化学者提出的网络时代会形成大量平行的文化提出了自己的质疑。作者认为未来的受众的中心集中度仍然会很高,不应该迷恋所谓的长尾,受众的分离,赢者通吃(winner-take-all)仍然是主要规律。^[29]当然,如果增加了前面对于研究对象的背景性考虑之后,研究就会更加深入,例如后期的学者又针对这个话题进行了跨国家的研究,结果发现更严重的碎片应该发生在不同的小的、新兴的媒体之间,而大媒体的受众重合度还是非常高的。^[30]

可以看到,大数据的挖掘体现出的属性差异还相对有限,都是以国家、语言、文化等差异为主,分析发现的维度并不多,缺乏更丰富的阐释深度。

(二)整体网中的属性自变量

在整体网中,属性数据也有不少直接做类似自变量的,但针对这种“点属性”,基本上进行的都是

描述性的说明,没有更多的量化方法直接支持关于点属性方面的判断。

如学者研究了国际音乐贸易的相关情况,通过 2002—2006 年 5 年的数据分析发现:美国、欧洲国家,包括德国、英国、荷兰处在贸易的中心地位。从五年的时段分析来看,这种结构非常稳定^[31]。这里,研究者其实是直接通过对数据的总结,得到了中心节点大多来自欧洲国家这个属性化的总结。用同样的理念,有研究分析了世界互联网的物理连接状态,8 年间的网络密度从 0.03 增加到了 0.034, MENA(中东北非)国家在 8 年中有不少进入了网络的中心圈,包括阿联酋、卡塔尔、埃及以及沙特阿拉伯,但是整体来看发展仍然符合强者恒强的规律。另一个有趣的发现在于,政治运动会通过互联网进行传播,阿拉伯之春是在中东北非的网络连通进入中心圈之后迅速产生的^[32]。可见,对国家的地理属性的分析,欧洲、中东、北非等是这类研究的主要发现维度,而且这种发现相对简单,只是对网络构成的一种现象性的总结。

还有一类研究虽然对点属性进行了更为量化的分析,但是采用的分析工具是传统的 SPSS 对均值差异进行显著性检验。

例如,有学者发现女性学者相较于男性学者只能获得更少的学术引用,这种性别的差距在学术引用时依然非常明显。作者首先通过引用网络分析得到了个体被引以及引用的相关数据,然后对包含性别在内的各项类目进行多维度的 ANOVA 分析得到了上述结论。作者采用了社会网的视角来解读,之所以出现这一现象是因为女性较难构筑自己的学术共同体,相比较而言,男性更倾向于与同性别的人进行学术合作。^[33]如果从个体具有较强相关性的角度来看这个数据分析,可能就会有一定的质疑。

当然在全网数据环境下,也出现了使用 ERGM 等方法进行分析的结构性研究。如有研究关注美国国会议员在 Twitter 上的关注网与互动网,结果发现关注网呈现了很强的党派同质性,影响互动网同质性的因素包括是否属于同一个州、党派、大众关注(public concern),如果控制了网络的内生变量之后,这种同质性程度会降低,所有的连接,包括关注连接和互动连接的存在本身都会增加投同质群体票的几率。^[34]

(三) 语义网络

语义网络的相关研究是社会网络研究当中不可缺少的一部分,也被认为是大数据环境下开始取代内容分析的一种方式。^[35]但是单纯的语义研究与本文的相关度不大。例如在研究埃及革命的语义表达时,revolution 一词的中心度最高,而 protest 则非常边缘,体现了大家对此事件的定性^[36];或者几个中国学者研究中国人现阶段的隐私观念,用新浪微博上 18000 条包含“隐私”一词微博构建一个词汇“共现”语义网络,然后通过聚类分析(cluster analysis)发现 11 个不同的语义网络等等^[37]。因为语义本身具有自己的含义,直接对其分析即可,因此多数都从框架理论或者词义角度进行分析,也可以从议程设置理论等视角直接进行分析^[38],这与属性数据的研究并没有关系。

但是如果语义网络与使用者网络结合进行分析,使用者属性就成为影响语义网络的重要要素,使用者属性成为自变量。

如学者讨论美国 Twitter 上群体互动的模式(pattern),主要进行两类研究:第一,在预先选定的几个关键词之下,考察用户(user)之间相互转发而形成的彼此连接情况,会发现用户大体分为多少个群组(cluster),这是话题讨论主题的聚类分析。第二,在群组之内,分析每条信息(message)的政治倾向,进行内容编码;同时看每条信息中提供了哪些链接,对链接的内容进行编码。这部分是话题立场分析。每个话题都选择了最近对此话题发表看法的 500 个使用者(node)。最后发现:从群组的角度看,极化明显,不同的群组中,自由派与保守派的比例呈现显著差异。而且圈子当中意见表达的信息条数数量越多(无论左右,加总的总量越多),则自由派与保守派比例相差越悬殊,也就是说群体极化越严重。^[39]

随后 Himelboim 等考察了每个话题网络中的情感取向问题。他们发现一个网络当中有一种占主导的情感。整个讨论大致形成正向、中心和负面三种情感分别主导的圈落。而每个圈子的情感是哪一种属性主要与其话题的内容、话题的意识形态有关,自由主义有更多的肯定性情绪,而保守主义则更多的是否定性情绪。该文是同质性的一次扩展,作者认为可以说发现了“情绪回音壁”(affective echo chambers)。^[40]

五、我国的 SNA 研究

前面分析的研究都是来自国外传播学顶级期刊中发表的文章,它们在方法的传承上有较为明显的脉络,发展相对连续、有逻辑。本部分则主要考察一下国内的相关研究状况。

与国外 SNA 相对清晰的脉络不同,我国传播学中的 SNA 研究一开始就体现出对“关系研究”的极大兴趣,尤其是与整体网结构相关的内容,诸如网络中心势、各类中心度、网络密度、网络子群等等。主要原因是我国开始较为丰富的 SNA 研究时就已经是社交网络相对较为盛行的年代,学者亟须对社交网络进行整体化、结构化的分析。因此,相比国外延续社会学传统,关注个体网中属性数据的重要性不同,我国的属性研究是非常有限的。比较典型的,如罗昕等在对广州的两个跑群的分析中,发现交流环境、组织管理、专业指导、社会资本、忠诚度等因素是影响虚拟社群信任网络生成的重要因素。^[41]韦路等研究了各国的媒介机构在 Twitter 上的网络构建情况,发现除了所在大洲的发达程度外,社会化媒体使用程度、地理接近性和文化接近性等国家因素都能够显著预测媒介机构之间的双向链接关系。^[42]

在大量的 SNA 研究中,成果最为丰富的领域是在知识图谱方面。主要体现出学者们期待能够更系统地了解国外在某类研究方面的现状,从而为具体话题的研究者提供一个全景性的认知导引,减少因为语言在阅读和了解文献上面的早期障碍,这部分研究大体都是纯粹的结构化分析,加上一些描述性内容,基本没有与属性变量结合进行复杂量化分析的案例。例如刘毅对国外舆论学研究的“知识图景”的研究^[43],包洪岩等对健康传播学的知识图谱的可视化研究^[44],陈艳红等对国外微博研究热点、趋势及研究方法的分析^[45],丁汉青等对境外传媒经济研究热点与场域的探索^[46]等等。

当然国内关于网络结构的各项指标的研究也会涉及各种热点话题,比如近期对于“一带一路”高峰论坛在 Twitter 上传播的子群结构的研究^[47]。

值得注意的是,我国已经有青年学者开始关注属性与关系之间的竞争关系,这体现出年轻学者与国际学术研究正在逐渐接轨。张伦在分析社交网络中的好友关系构建时,考察了个体特征、对偶特征以及网络结构特征对在线关系构建的影响,结果发现,网络结构特征相比于其他特征对于个体在线关系的构建更为重要^[48],这一结论与第三部分 Mai 等人的研究^[15]是非常一致的,即发现了既有的结构性因素其实超越了个体的属性因素,成为网络关系新连接的核心影响因素。

六、结语:关系与属性数据的再结合

Shumate 等曾经把社会网络分析的诸多研究依据他们之间形成关系内容的不同,划分为四种类型:亲近关系(affinity)、信息流(flow)、所属关系(representational)以及语义网络(semantic)。^[35]他们认为依据“关系”进行分类的方法适用于人际、组织、大众、健康、政治传播以及计算机中介传播(CMC),打破了传统的依据点状态区分为人际、组织、群体、大众的习惯性思维。从这种分类的思路中可以看出社会网络学者们期待从社会网络的“关系”视角重新审视和思考传播学的诸多子领域的研究。

而 Cappella 在 2017 年也发表了一篇文章^[49],谈大众传播与人际传播未来的研究方向,他提到其中的一个方向就是未来需要融合人际传播与大众传播,融合的方式就是从因素(factor)的研究,或者

说是变量(variable)研究过度到个体(actor)的研究。他也强调,以前对于变量或者说因素的研究前提是作为变量载体的人是彼此独立的,但是如果人是彼此联系的,可能就需要转换思路,把研究的核心换为个体(actor),这是社交网络普及的今天所需要的。基本的理念就是个体作为多个变量的携带者(carrier),以前是K个独立的个体拥有N个变量的模型,而现在则是拥有一个 $N * K$ 变量的模型,可以认为变量是一种对大众传播过程的描述,而个体则体现了对人际传播结构的重视。代理人基模型(agent-based modeling, ABM),当然也包括多层次模型(multilevel models)可以实现上述目标,作者非常期待传播学者能够运用各种新的研究工具,同时考量个体与变量两类内容。

以ABM为例,该模型的主要理念是把现实抽象为个体(agent)和环境(environment),个体可以是任何具有主动性的主体,包括个人、公司、国家等,然后把个人与环境、个人与个人的互动抽象为可表达的数学公式。比如,在群体暴力模型中可以抽象出的个体为激进分子、平民和警察,三者都会与环境互动,三者之间彼此也会互动。如果平民会产生暴力行为的影响因素之一是其群体具有的愤怒情绪,而该情绪的愤怒程度可以抽象为:专制性*(1-合法性),即主体对政府的专制程度以及合法性的态度的函数,可以代表该主体的愤怒指数。当愤怒指数到达角色转变点时,平民会发生角色转变,成为激进分子;类似的,还可以构建多个模型,包括主体被警察抓到的概率等等,当概率超过一定水平时,主体会放弃转变为激进分子。当所有的变量间以及变量和环境间模型构筑好之后,通过仿真模型再现群体暴力的场景,进而从定量的角度分析产生暴力行为的动力学特征。^[50]

Cappella从大众传播与人际传播结合的角度赋予了“关系”变量与“属性”变量结合的一个更广泛,也更具理论内涵的意义。随着新的发展,社会网络研究的学者应该也会在个体与属性的结合上继续进行探讨,而不仅仅局限在现在使用较多的ERGM和SIENA模型,后两者的研究虽然基于关系数据与属性数据的结合处理,但是最终解决的问题相对单一,即两个点之间构建连接的概率是否大于随机网络下构建连接的概率,侧重点仍然在于网络的构建,未来应该期待更为丰富的方法与理论的支持。

参考文献:

- [1] L. Freeman. The Development of Social Network Analysis. Vancouver; Empirical Press, 2004; 1.
- [2] V. D. Miller, M. S. Poole, D. R. Seibold, et al. Advancing Research in Organizational Communication Through Quantitative Methodology. Management Communication Quarterly, 2011, 25(1): 4-58.
- [3] Y. Kim, S.-H. Hsu, H. G. de Zúñiga. Influence of Social Media Use on Discussion Network Heterogeneity and Civic Engagement: The Moderating Role of Personality Traits. Journal of Communication, 2013, 63(3): 498-516.
- [4] S. W. Campbell, N. Kwak. Mobile Communication and Strong Network Ties: Shrinking or Expanding Spheres of Public Discourse? New Media & Society, 2012, 14(2): 262-280.
- [5] G. A. Barnett, G. A. Benefield. Predicting International Facebook Ties through Cultural Homophily and other Factors. New Media & Society, 2016, 19(2): 217-239.
- [6] Y. Arie, G. S. Mesch. The Spatial and Social Network Dimensions of Mobile Communication: A Test of the Social Stratification and Social Diversification Hypotheses. Communication Research, 2016, 43(5): 713-734.
- [7] H. Farrow, Y. C. Yuan. Building Stronger Ties with Alumni Through Facebook to Increase Volunteerism and Charitable Giving. Journal of Computer-Mediated Communication, 2011, 16(3): 445-464.
- [8] T. Kobayashi. Bridging Social Capital in Online Communities: Heterogeneity and Social Tolerance of Online Game Players in Japan. Human Communication Research, 2010, 36(4): 546-569.
- [9] C. Shen, P. Monge, D. Williams. Virtual Brokerage and Closure: Network Structure and Social Capital in a Massively Multiplayer Online Game. Communication Research, 2014, 41(4): 459-480.
- [10] H. Lee, N. Kwak, S. W. Campbell. Hearing the Other Side Revisited: The Joint Workings of Cross-Cutting Discussion and Strong Tie Homogeneity in Facilitating Deliberative and Participatory Democracy. Communication Research, 2015, 42

- (4):569-596.
- [11] J. K. Lee, J. Choi, C. Kim, et al. Social Media, Network Heterogeneity, and Opinion Polarization. *Journal of Communication*, 2014, 64(4):702-722.
- [12] P. Monge, S. Lee, J. Fulk, et al. Research Methods for Studying Evolutionary and Ecological Processes in Organizational Communication. *Management Communication Quarterly*, 2011, 25(2):211-251.
- [13] M. L. Doerfel, M. Taylor. The Story of Collective Action: The Emergence of Ideological Leaders, Collective Action Network Leaders, and Cross-Sector Network Partners in Civil Society. *Journal of Communication*, 2017, 67(6):920-943.
- [14] 彼得·芒戈, 诺什·康特拉克特. 传播网络理论. 陈禹, 刘颖译. 北京: 中国人民大学出版社, 2009:44;51.
- [15] B. Mai, J. Liu, S. González-Bailón. Network Effects in the Academic Market: Mechanisms for Hiring and Placing PhDs in Communication(2007-2014). *Journal of Communication*, 2015, 65(3):558-583.
- [16] M. Shumate. The Evolution of the HIV/AIDS NGO Hyperlink Network. *Journal of Computer-Mediated Communication*, 2012, 17(2):120-134.
- [17] D. B. Margolin, C. Shen, S. Lee, et al. Normative Influences on Network Structure in the Evolution of the Children's Rights NGO Network, 1977-2004. *Communication Research*, 2015, 42(1):30-59.
- [18] G. A. Barnett, J. A. Danowski, T. H. Feeley, et al. Measuring Quality in Communication Doctoral Education Using Network Analysis of Faculty-Hiring Patterns. *Journal of Communication*, 2010, 60(2):388-411.
- [19] H. Song. Uncovering the Structural Underpinnings of Political Discussion Networks: Evidence from an Exponential Random Graph Model. *Journal of Communication*, 2015, 65(1):146-169.
- [20] W. Lu, K. N. Hampton. Beyond the Power of Networks: Differentiating Network Structure from Social Media Affordances for Perceived Social Support. *New Media & Society*, 2016, 19(6):861-879.
- [21] B. Stephens, W. Chen, J. S. Butler. Bubbling Up the Good Ideas: A Two-Mode Network Analysis of an Intra-Organizational Idea Challenge. *Journal of Computer-Mediated Communication*, 2016, 21(3):210-229.
- [22] B. Etling, J. Kelly, R. Faris, et al. Mapping the Arabic Blogosphere: Politics and Dissent Online. *New Media & Society*, 2010, 12(8):1225-1243.
- [23] S. Vicari. Exploring the Cuban Blogosphere: Discourse Networks and Informal Politics. *New Media & Society*. 2015, 17(9):1492-1512.
- [24] Y. S. Lim, H. W. Park. The Structural Relationship between Politicians' Web Visibility and Political Finance Networks: A Case Study of South Korea's National Assembly Members. *New Media & Society*, 2013, 15(1):93-108.
- [25] C. Woo-young, H. W. Park. The Network Structure of the Korean Blogosphere. *Journal of Computer-Mediated Communication*, 2012, 17(2):216-230.
- [26] S. Dvir-Gvirsman, Y. Tsfati, E. Menchen-Trevino. The Extent and Nature of Ideological Selective Exposure Online: Combining Survey Responses with Actual Web Log Data from the 2013 Israeli Elections. *New Media & Society*. 2016, 18(5):857-877.
- [27] J. Shin, L. Jian, K. Driscoll, et al. Political Rumoring on Twitter during the 2012 US Presidential Election: Rumor Diffusion and Correction. *New Media & Society*, 2016, 19(8):1214-1235.
- [28] H. Taneja. Mapping an Audience-centric World Wide Web: A Departure from Hyperlink Analysis. *New Media & Society*, 2016, 19(9):1331-1348.
- [29] J. G. Webster, T. B. Ksiazek. The Dynamics of Audience Fragmentation: Public Attention in an Age of Digital Media. *Journal of Communication*, 2012, 62(1):39-56.
- [30] R. Fletcher, R. K. Nielsen. Are News Audiences Increasingly Fragmented? A Cross-National Comparative Analysis of Cross-Platform News Audience Fragmentation and Duplication. *Journal of Communication*, 2017, 67(4):476-498.
- [31] S. -I. Moon, G. A. Barnett, Lim Yon Soo. The Structure of International Music Flows Using Network Analysis. *New Media & Society*, 2010, 12(3):379-399.
- [32] H. Seo, S. J. Thorson. Networks of Networks: Changing Patterns in Country Bandwidth and Centrality in Global Information Infrastructure, 2002—2010. *Journal of Communication*, 2012, 62(2):345-358.

- [33] S. K. Westerwick, C. J. Glynn. The Matilda Effect—Role Congruity Effects on Scholarly Communication: A Citation Analysis of Communication Research and Journal of Communication Articles. *Communication Research*, 2013, 40(1): 3-26.
- [34] T. -Q. Peng, M. Liu, Y. Wu, et al. Follower-Followee Network, Communication Networks, and Vote Agreement of the U. S. Members of Congress. *Communication Research*, 2016, 43(7): 996-1024.
- [35] M. Shumate, A. Pilny, Y. Catouba, et al. A Taxonomy of Communication Networks. *Annals of the International Communication Association*, 2013, 37(1): 95-123.
- [36] Z. Papacharissi, M. de Fatima Oliveira. Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt. *Journal of Communication*, 2012, 62(2): 266-282.
- [37] E. J. Yuan, M. Feng, J. A. Danowski. "Privacy" in Semantic Networks on Chinese Social Media: The Case of Sina Weibo. *Journal of Communication*, 2013, 63(6): 1011-1031.
- [38] Z. Tai. The Structure of Knowledge and Dynamics of Scholarly Communication in Agenda Setting Research, 1995-2005. *Journal of Communication*, 2009, 59(3): 481-513.
- [39] I. Himelboim, S. McCreery, M. Smith. Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter. *Journal of Computer-Mediated Communication*, 2013, 18(2): 40-60.
- [40] I. Himelboim, K. D. Sweetser, S. F. Tinkham, et al. Valence-based Homophily on Twitter: Network Analysis of Emotions and Political Talk in the 2012 Presidential Election. *New Media & Society*. 2016, 18(7): 1382-1400.
- [41] 罗昕, 许倩婷. 虚拟社群信任网络的生成与维系: 以广州两个跑群为例. *国际新闻界*, 2014, 9: 58-73.
- [42] 韦路, 丁方舟. 社会化媒体时代的全球传播图景: 基于 Twitter 媒介机构账号的社会网络分析. *浙江大学学报(人文社会科学版)*, 2015, 6: 91-105.
- [43] 刘毅. 国外舆论学研究的“外知识图景”: 热点、网络与结构——基于 SSCI 数据库(1994-2013)的知识图谱分析. *新闻与传播研究*, 2015, 5: 19-31.
- [44] 包洪岩, 殷晓莉, 兰小筠. 基于知识图谱的健康传播学可视化研究. *新闻与传播研究*, 2013, 12: 107-118.
- [45] 陈艳红, 宗乾进, 袁勤俭. 国外微博研究热点、趋势及研究方法: 基于信息计量学的视角. *国际新闻界*, 2013, 9: 28-40.
- [46] 丁汉青, 曹璞. 2013 年境外传媒经济研究热点与场域——基于文献计量学的方法探索. *新闻与传播研究*, 2015, 4: 61-82.
- [47] 汤景泰, 星辰, 高敬文. 论“一带一路”国际话语权的提升——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*, 2018, 5: 56-65.
- [48] 张伦. 个体在线网络关系构建影响因素研究. *国际新闻界*, 2017, 4: 25-43.
- [49] J. N. Cappella. Vectors into the Future of Mass and Interpersonal Communication Research: Big Data, Social Media, and Computational Social Science. *Human Communication Research*, 2017, 43(4): 545-558.
- [50] 陈鹏, 陈建国, 袁宏永. 群体性暴力行为的多主体建模研究综述与思考. *系统仿真学报*, 2015, 3: 439-447.

Rethinking the Relational Data and Attribute Data: The Status and Development of Social Network Analysis

Pang Yunxia (Central University of Finance and Economics)

Abstract: The advantage of the social network analysis(SNA) lies in the analyzing ability of the "relational" data, which overcome the shortcomings of traditional variable analysis for "attribute" data. However, lack of the ability to analyze the "attribute" data may also be SNA's weakness. With the emphasis of organizational communication scholars on the multi-level analysis of "relational" data and "attribute" data, this paper examines the application of SNA in various research fields of communication in recent years from the per-

spective of how they treat "attribute data" and "relational data" in different ways. This article finds that early SNA researches focused on individual networks, which was from the tradition of sociology, attribute data and relational data are usually analyzed separately, either could be independent or dependent variables. However there could be a premise error of such individual network analysis, because variable analysis demands individuals are independent with each other, which just can not be accorded in a personal network. Then organizational communication scholars initiated ERGM and SIENA models, and began to put these two types of data in the same model. The advantage of these new models are that they can deal with data from agents not independent with each, and can do comparison of the effects from endogenous (network) and exogenous (attribute) variables. Recent researches control one kind of variables from above and find that endogenous variables, i. e. network factors, are the main dynamics of tie formation instead of exogenous (attribute) factors. In the network environment, researchers use data from online and analyze from the perspective of whole network of SNA. Results coming out of data digging just opposite to traditional research methods with hypothesis put forward in advance. They often find culture, country, language and location are the key factors forming clusters. That means online social network analysis step back to analyze relational and attribute data separately. However, there are also researchers now also trying to combine these two kinds of data in one research. In general, SNA begins to emphasize the joint exploration of relational data and attribute data in recent years. However, although the research methods are developed to analyze these two types of data, the effect of this combination is not ideal, and more work needs to be done in the future, such as the emergence of new research tools and theories suitable for this trend.

Key Words: social network analysis; research progress; attribute data; relational data

■收稿日期:2018-11-25

■作者单位:庞云黠,中央财经大学文化与传媒学院;北京 100081

■责任编辑:汪晓清