

水浒人物社会网络统计分析^{*}

钮亮

(中国计量大学经济与管理学院, 杭州 310018)

摘要: [目的/意义] 揭示水浒人物排名和人际关系形成的原因。[方法/过程] 借助多元线性回归模型分析水浒网络中心性对座次排名的影响, 通过指数随机图模型分析水浒人际关系的成因。[结果/结论] 在排名分析中发现度中心性和介数中心性对排名有正向影响关系; 在人际关系形成分析中发现, 地理接近性、官僚、贫民和富商对网络形成影响显著, 网络结构中度中心性和接近中心性对人物关系的形成存在正向影响。

关键词: 水浒成员网络 斯皮尔曼等级相关 回归模型 指数随机图模型

分类号: TP391

DOI: 10.31193/SSAP.J.ISSN.2096-6695.2020.02.10

0 引言

《水浒传》是中国四大名著之一, 历来研究者不乏其人。水浒研究大致可分为版本、成书探索^[1]、海外翻译^[2]、思想研究^[3]、性格探索^[4]等几种类型, 这些研究对推动理解这部名著做出了贡献。但他们采用的是传统人文研究方法, 通过思辨、批判、创作、勾画、想象、阅读、提炼等完成对作品的理解, 想象多于实证。人文研究可以得出富有创见性的结论, 但这个结论不以正确或错误为判准, 也不能诉诸证伪的方式来得到澄清, 一般属于情感、道德和精神世界, 其作用是让心灵具有越来越深邃的体验。而定量研究依靠实验、数据、图表、模型、证实、证伪、复制等方法达到对作品的理解^[5], 能够捕捉事情的前因后果, 实现对作品中信息的重复验证。尽管传统人文研究也会对事情的发展给出因果判断, 但大多是依据现成的理论, 通过演绎推导来理解和解释作品的。定量研究方法以归纳为导向, 并不预设某个既定理论的存在, 一般通过统计手段达到对作品的解释。如果找到一个能够对被解释事件的概率造成影响的因素, 那就有了解释^[6]。这种解释实现了波普尔说的证伪, 将对作品的理解导向科学。

^{*} 本文系国家社会科学基金项目“我国标准‘走出去’与贸易强国战略研究”(项目编号: 18BJL102)的研究成果之一。

【作者简介】钮亮(ORCID: 0000-0001-6934-4416), 男, 副教授, 研究方向为数字人文、知识创新, Email: niutyut@126.com。

具体而言,本研究主要聚焦于两个问题:(1)探究水浒人物座次排名的影响因素;(2)探究水浒人物关系网络形成的原因。

1 水浒人物社会网络建构

为了构建水浒人物的社会网络,需要从水浒文本中提取人物之间的关系,构建社会网络,实现从非结构化信息到结构化信息的转化。由于中文语句中词项之间没有空格,人物的关系匹配首先需要进行中文分词。现有的分词算法可分为三大类:基于字符串匹配的分词方法、基于理解的分词方法和基于统计的分词方法。基于字符串匹配的分词方法是按照一定的策略将待分析的汉字串与词典中的词条进行匹配,若在词典中找到某个字符串,则匹配成功。按照扫描方向的不同,基于字符串匹配的分词方法可以分为正向最大匹配算法和逆向最大匹配算法,二者原理一致,只是前者是从字符串前面计算,后者是从字符串后面计算。本文人物社会网络构建受赵京胜等人研究方法的启发^[7],采用字符串匹配的分词方法,事先建立水浒人物词典,将水浒文本中人物的名字、字号、浑名做同义归并,例如将文中宋江、宋公明、及时雨的名称统一替换为宋江。然后,将待分析的文本与这个词典进行匹配,按照从左向右的顺序对小说进行词库匹配切分,当识别出它们在一个段落时就两两建立关系,出现一个人物的则作为孤立点。由于本文重点分析水浒人物社会网络的成因,对于两个人物多次出现的关系按照一次计算,不考虑加权网络情况。构建的网络如图1所示:

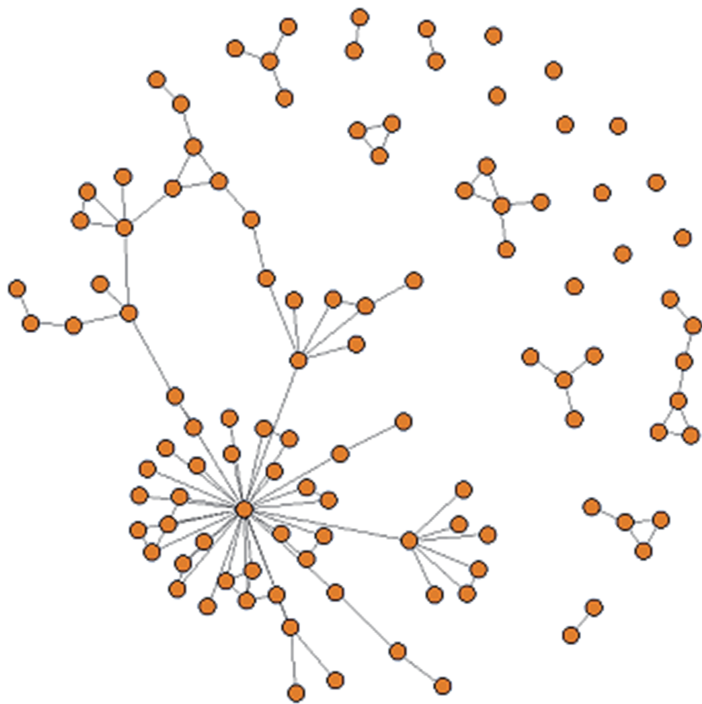


图1 水浒人物社会网络

2 水浒排名的影响分析

梁山聚会后七十一回提到“忠义堂石碣受天文，梁山泊英雄排座次”的事情。按照书中的说法，这个座次是神授如此。但是按照唯物论的立场，人的本质是由社会关系的综合决定的。就排名来说，笔者想考察这个座次的建立是否和梁山好汉在网络中的位置有关。从直观上判断，处于独特网络位置的人，其排名也应该比较高。本研究尝试通过定量分析来证实这个判断。

社会网络位置的测量方法具体包括度中心性、介数中心性、接近中心性、特征中心性、结构洞等^[8]，是衡量人物社会资本的重要参量^[9]，反映人在社会网络结构中的重要性、优越性或特权性。网络节点的各种中心性决定了节点控制和掌握社会资源的能力，这些能力可能是排名大小的影响因素。因此，可以通过构建以网络中心性为自变量，以排名为因变量的回归模型来测度排名受中心性影响的程度。以网络中心性作为自变量构建回归模型不乏先例，已有文献在自变量的选择方面主要集中在度中心性^[10]、介数中心性^[11]、接近中心性^[12]、结构洞^[13]等单一方面。由于每种中心性所表达的社会资本内涵不同，在解释水浒人物排名上起到的作用也不同，因此本文选择度中心性、介数中心性、接近中心性、结构洞 4 种中心性作为自变量进行模型构建。由于数据归一化后，最优解的寻优过程会变得平缓，更容易正确的收敛到最优解，而且得出的系数也更有解释力^[14]，因此，本研究将所选的因变量和自变量都做归一化处理。依照每种中心性表达的内涵，以及它们在揭示排名中的效果，本研究设定四个假设，通过建立多元线性方程来检验这些假设成立与否。

H1a: 人物的度中心性越大，则其排名越高；

H2a: 人物的介数中心性越大，则其排名越高；

H3a: 人物的接近中心性越高，则其排名越高；

H4a: 人物的结构洞越大，则其排名越高。

模型计算结果如表 1 所示。

表 1 排名与中心性的线性模型分析

中心性作为自变量	估计值	标准误	P 值	R ²	F
Degree	5.133e-01	1.295e-0	0.000137***	0.7165	65.07
Between	3.616e-01	1.348e-01	0.008514**		
Closen	-8.284e-02	5.455e-02	0.131915		
Burt	-1.309e-02	5.777e-02	0.821191		

注: ***表示 $p < 0.001$, **表示 $p < 0.01$, *表示 $p < 0.05$

R² 的值为 0.7165，说明当前模型能够解释水浒人物排名变异情况的 71.65%。F 统计量反映了这些变量中至少有一个可以解释排名情况，验证了模型的有效性。

从回归分析结果可以发现,对人物排名具有显著性影响的变量是度中心性和介数中心性,证实了 H1a、H2a 的假设要求,其他两项假设由于接近中心性和结构洞的 p 值不明显,不具有解释效力。部分人物的度中心性、介数中心性和排名如表 2 所示。

表 2 部分人物的度中心性、介数中心性和排名

姓名	度中心性	介数中心性	排名
宋江	33	1832.5	1
卢俊义	5	332.5	2
吴用	2	126	3
公孙胜	1	0	4

就 H1a 和 H2a 来说,宋江是典型代表。宋江绰号为“及时雨”,结交了诸多英雄好汉,他的度中心性达到 33,而且他在江湖和官府这两种异质人群都有人认识,通过他可以认识不同的成员,他的介数中心性达到 1832.5,可见他在水浒排行中位列第一不是没有道理的。其中有争议的是公孙胜排名第四,他的度中心性和介数中心性都不高。这说明,尽管度中心和介数中心性对排名有一定影响,但还存在其他影响因素,比如人物的身份、资历等。公孙胜与晁盖、吴用等七人结义,一同劫取生辰纲,后上梁山入伙,是梁山掌管机密的人物,其资历和身份因素对排名影响更大。

3 水浒社会关系形成的网络统计分析

一般社会网络的研究都是描述性的,不关心网络生成的原因,但是对网络产生的原因分析既可以帮助了解现有网络的形成机制,也为未来新网络的生成提供了解释可能。按照波普尔的说法,“衡量一种理论的科学地位的标准是它的可证伪性或可反驳性或可检验性”^[15]。水浒人物网络为什么会形成以宋江为首的网络,而不是一个随机网络?有哪些原因导致它产生了这个结果?能否通过仿真再现网络的形成?如果再现是可行的,那就说明本研究具有可重复、可验证性。

3.1 水浒观察网是随机生成的吗?

为了寻找影响网络形成的原因,首先需要明确观察网络不是随机网络,如果网络是随机的,那么其形成就不遵循某种机制,分析它的形成原因也就无意义了。一个观察网络的构成可以分解为边共享、二组共享(DSP)、二组边共享(ESP)和三角形、 k 星单元等,其中最小的单元是边,边共享、二组共享和三角形反映了网络的集聚程度。要探测一个观察网络是否由某种规则生成而不是随机生成的,就是选取边数量和密度都相同的观察网络和随机网络进行构成单元比较,看这些构成单元的数量和随机网络是否不同。计算参考 Jenine K.Harris (2013) 研究的成果,通过 R 语言和 statnet 包实现^[16],结果见图 2 所示。与同样边数和密度的随机网络相比,水浒传

络的度中心性、二元组共享、二元组边共享都比随机网络表现优异, 说明水浒网络的生成是遵循一定规则的。

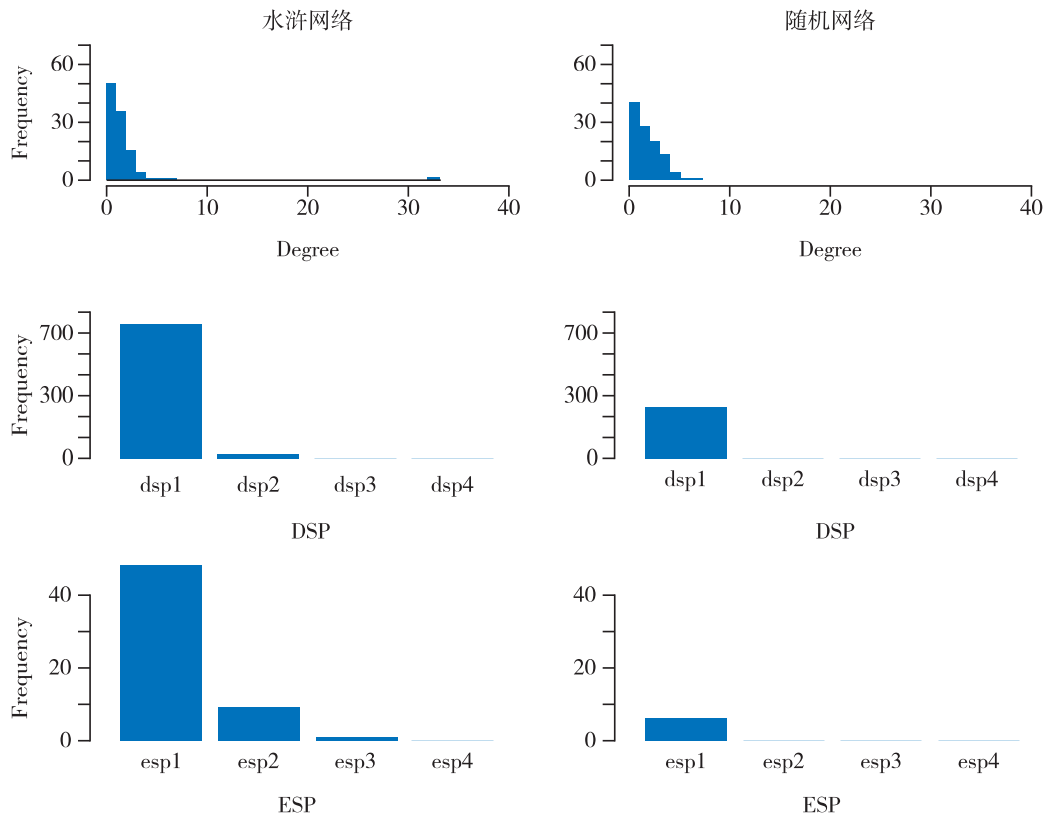


图 2 水浒网络与随机网络比较

如果网络由某些规则构成, 那么它更容易产生一种聚合效应。在网络的节点关系中, 如果两个人有一个共同的朋友, 则这两个人在未来成为朋友的可能性就会提高。这两人成为朋友后, 他们之间就构成了三角形。如果一个网络中存在着超出随机网络的三角形数量, 也说明网络的生成是遵循一定规则的。本研究对 100 个和水浒人物网络同样边数和密度的随机网络进行了仿真计算, 并和水浒人物网络进行比较, 结果发现, 水浒人物网络的三角形数量达到 23 个之多, 而 100 个随机仿真网络中三角形数量却很少。大多数仿真网络的三角形数量为 0 个, 其中有一个仿真网络的三角形数量为 7 个, 是拥有三角形数量最多的网络, 但即便如此也没法和水浒网络的三角形数量相比。这证明了水浒人物网络的生成是遵循某种规则的。随机网络和水浒网络中的三角形数量见图 3 所示, 其中纵坐标为网络数量, 横坐标为网络中对应的三角形数量, 菱形块所在位置是水浒网络对应的三角形数量。

3.2 梁山好汉网络形成机制

3.2.1 分析原理

通过图 2、图 3 的分析, 大致可以判定水浒人物网络的形成是受某些规则支配的, 它或者由

网络子图的某种拓扑规则自动转换形成,例如“出二星(out-2-star)”就倾向于形成三角形;或由网络成员的个体特征、他们在交互项规则中形成的共同属性(例如成员若处于同一地区,则地区属性就是他们的共同属性)^[17]等节点属性之类的外生协变量触发形成。

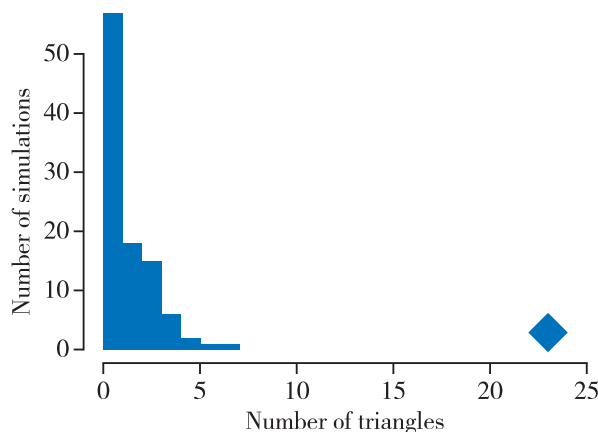


图3 水浒传的三角形与100个随机网的三角形

基于拓扑规则对网络形成的影响在社会学中被称为社会影响模型,致力于发展这种理论的有齐美尔、布劳、怀特、科尔曼、林南、格拉诺维特、博特诸人^[18]。基于外生协变量对网络形成的影响在社会学上被称为社会选择模型,研究节点属性对网络形成的贡献。网络的形成具有趋同性,指网络关系倾向于在具有相同属性的节点之间发生。二元关系协变量可以用来表示人们是否具有相同的属性或成员关系,影响到网络关系的生成。受《社会网络指数随机图模型》^[19]一书中社会选择模型案例的启发,本文采用社会选择模型来理解水浒人物集聚的成因。

本研究初步选定的节点属性为身份、地理区域、社会资本(度中心性、介数中心性、接近中心性、结构洞),并通过指数随机图模型(exponential random graph model,简称ERGM)考察这些属性对网络形成的影响。指数随机图模型是以关系为基础的模型,它把社会理论转换成以关系形式表达的假设,然后通过统计运算检验假设是否成立^[20]。

3.2.2 设定假设

为了研究人物属性变量对网络形成的影响,特设定以下几种假设:

H1b: 相似身份的水浒人物容易建立交往关系。

同一阶层在习惯、教养、价值观上更容易一致,根据“物以类聚,人以群分”规则,他们倾向于彼此结合。按照水浒人物上梁山之前的出身,大致分四种类型,分别是贫民、商贩、地主和官僚,并对其进行相应的编码,贫民编码为1,商贩编码为2,地主编码为3,官僚编码为4。在这四种身份中假设相似身份的人物容易建立交往关系。

H2b: 水浒人物之间地理接近性越高,越有可能建立交往关系。

Daraganova等人认为,地理接近性越高的人员,越容易建立联系^[21]。若两个水浒人物在同

一省份, 就认为具有较高的地理接近性。例如: 宋江与吴用同属于山东省, 按照假设, 他们应该更容易建立联系。

针对水浒人物的社会结构资本, 提出了下面四种假设:

H3b: 度中心性越大的人越容易建立联系。

度中心性大, 说明连接他的人员越多, 他就越容易和别人形成连边关系。

H4b: 介数中心性越大的人和别人建立联系的可能性越大。

介数中心性是网络中所有最短路径中经过该节点的路径数目占最短路径总数的比例, 衡量行动者对资源信息的控制程度, 反映了节点作为“桥梁”的重要程度。

H5b: 结构洞越大的人越容易与他人建立联系, 反之亦然。

人总是信任特定的人, 并依赖于特定的人产生交换, 当双方关系并不十分密切时, 与双方关系都很密切的第三个人就占据了一个结构洞, 连接结构洞的节点就成为桥梁。同质网络中, 很难产生结构洞。因为同质网络当中存在大量冗余的信息, 不同的集群之间往往因为具有不同的信息而存在沟通的必要。本研究运用博特 (Burt) 提出的网络约束系数来测度结构洞。网络约束系数综合考虑了节点度和节点邻居拓扑关系信息, 网络约束系数值越大, 说明该节点邻居数量越少且与其邻居间的闭合程度越高, 结构洞越小, 这样的节点由于不易获得新的关系资源使得它在竞争中处于不利地位。反之, 网络约束系数值越小, 结构洞形成的机会就越大, 越有利于获得新的关系资源, 容易与他人建立关系。^[22]

H6b: 接近中心性越大的人越容易与他人建立联系。

接近中心性度量的思想是, 如果一个行动者与许多其他行动者都很“接近”, 那么该节点处于网络中心位置。因为一个行动者可能通过直接或者间接的方式接触到网络中的许多人, 在这种情况下, 他就具有高的接近中心性, 即这位行动者能接近网络中大量的其他行动者。因此, 接近中心性大的人应该更容易和他人建立联系^[23]。

3.2.3 假设检验

为了审视水浒人物网络的形成过程, 本研究将简单随机图模型作为理解整个水浒网络模型的基准。简单随机图模型是网络成员之间的随机分配关系, 不考虑网络成员的属性信息以及任何影响关系形成的社会力量, 其密度和边数与水浒观察网络相同, 但成员之间的关系是独立且随机发生的。在简单随机图模型基础上增加节点属性和节点交互项来仿真水浒网络的生成过程, 观察它对随机网络的改进, 就是社会选择模型的机制。本文将简单随机图模型称为模型 1, 将增加了节点属性和交互项的选择模型称为模型 2, 在模型 2 的基础上增加了网络中心性的选择模型称为模型 3。经过计算后形成的结果见表 3。

表 3 水浒网络的三个模型

模型	参数	估计值	新边概率	P 值	AIC	BIC
模型 1	Edges	-3.93288	0.019	<1e-04 ***	1099	1106
模型 2	Edges	-4.5202	0.011	<1e-04 ***	1059	1099

续表

模型	参数	估计值	新边概率	P 值	AIC	BIC
	贫民	0.2432	1.275	0.5748		
	商贩	0.5825	1.790	0.0472*		
	地主	1.0847	2.958	0.1427		
	官僚	1.2048	3.336	<1e-04 ***		
	省份	1.0640	2.898	<1e-04 ***		
模型 3	Edges	-6.4557133	0.015	<1e-04 ***	899	966
	贫民	0.9723355	2.644	0.031626*		
	商贩	0.8235472	2.278	0.008725**		
	地主	1.1964639	3.308	0.111075		
	官僚	0.9317222	2.538	0.000546***		
	省份	0.9523298	2.592	<1e-04***		
	Degree	0.1796862	1.197	0.000925***		
	Between	-0.0013944	0.998	0.149906		
	Burt	-0.4181213	0.658	0.134883		
	Closeness	43.8734664	1.132e+19	0.001490**		

注：*** 表示 $p < 0.001$ ，** 表示 $p < 0.01$ ，* 表示 $p < 0.05$

对于模型 1 来说，p 值显著的参数是 edges；对于模型 2 来说，分别是 edges、商贩、官僚和人物所在省份。对于模型的整体解释能力，在增加了节点属性及其交互项后，模型 2 的 AIC 和 BIC 值由原来的 1099 和 1106 变成 1059 和 1099，说明模型解释能力得到了增强；在模型 2 的基础上增加网络的节点中心性作为主效应，形成模型 3，进行仿真后发现，AIC 和 BIC 值由 1059 和 1099 变成了 899 和 966，模型解释能力又得到了增强。因此，本文用模型 3 来考察上文提出的 6 个假设。

假设 H1b 中，贫民的连接概率达到 264.4%，说明穷人容易和穷人连接，这反映出当民不聊生的时候，贫民常常会响应起义的号召而参加革命，中国历代的农民起义印证了这一点；商贩和官僚都对网络形成有显著影响，水浒人物中，商贩出身者形成关系的概率为 227.8%，官僚出身者建立关系的概率达到 253.8%，这两种出身者是梁山人物的主要来源，其中官僚阶层在水浒网络形成中的贡献特别显著，p 值最小。比较各种出身人物的平均联系次数（如图 4 所示），官僚出身者之间的联系也最多，这或许是因为水浒人物网络中官僚出身者占的数量多的原因；地主观

念保守无法形成一个团结的、主导未来局势的阶层, 统计结果也不显著。

对于假设 H2b 来说, 省份的 p 值非常显著, 同一地理区域的人物之间建立联系的概率达到 259.2%, 这说明, 地理接近性对于关系的形成具有非常显著的正向影响。

网络中心性对关系的形成也具有很大作用。度中心性统计显著, 测度后的连接概率是 119.7%, 验证了假设 H3b 成立, 表示节点的度中心性越大, 越容易与他人建立连接。介数中心性和结构洞的统计检验不显著, 说明假设 H4b 和假设 H5b 不成立。接近中心性对网络的形成影响力巨大, 概率高达 $1.132e+19$, 这大概是由于水浒人物之间大多是直接或者间接通过朋友认识的, 是典型的熟人社会, 因此接近中心性在水浒人物网络形成中达到这么高的贡献系数也就可以理解了。

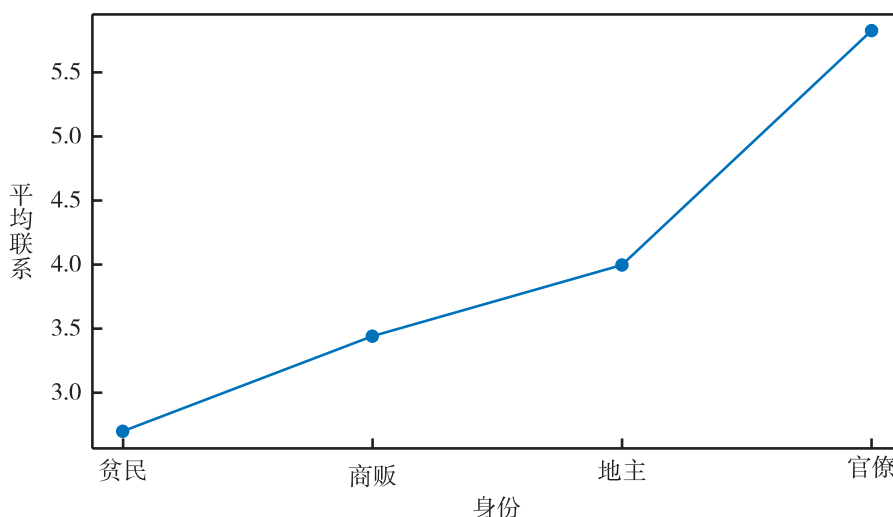


图 4 不同出身人物的平均联系次数

4 结 语

以往对水浒的研究多使用人文的定性方法, 很少用到定量分析的手段。从小说解释的角度来说, 基于统计立场的定量分析方法不事先设定理论, 以问题为导向, 通过找到可测变量来构建因果模型达到对问题的解答, 为人文研究的证实或证伪提供了可能^[24]。定量分析应用于古典文学研究, 其作用既可能是解决主观研究无力把握文学作品内容的局限, 更重要的是, 借助定量分析结果来提出文学研究的新问题, 促成定性研究和定量研究的结合, 推动文学研究向更广泛的领域发展。本文以定量方法分析水浒中的人物关系, 有以下发现:

(1) 通过分析水浒人物的座次排名同网络中心性(度中心性、介数中心性、接近中心性、结构洞)的关系, 借助多元线性模型考察中心性对排名的影响, 发现度中心性和介数中心性对水浒人物排名有正向影响;

(2) 通过社会选择模型计算人物的地理接近性、身份、网络中心性对水浒人物社会互动交往

关系的影响,发现人物所在省份、人物身份(主要是官僚、商贩和贫民)对水浒人物关系形成有正向影响;网络中心性中,度中心性和接近中心性对人物关系的形成有正向影响,其他中心性的统计特征不显著;

(3)在网络形成的统计分析中,通过AIC和BIC对模型的拟合优度进行评价,AIC和BIC值越低,说明仿真网络的拟合优度越好。后续研究可以选取更多的参数,扩充对人物属性的描述,论证其他属性对水浒网络形成的贡献,改进仿真网络的拟合优度,使得人文研究也具有了控制实验的特性。

总体而言,本研究立足于数字人文的方法论视角,基本立场是“以数据观数据”,通过数学统计和可视化技术呈现分析结果,弥补阅读者不能把握深层数据之间关系的缺陷。本研究可以作为分析中国古典小说人物关系网络的一个样例,为数字人文的实际应用提供参考。

【参考文献】

- [1] 温庆新.关于《水浒传》成书时间研究的方法论思考[J].清华大学学报(哲学社会科学版),2014(2):45-53.
- [2] 宋健飞.德译《水浒传》——欧洲最早、最奇、最经典的译本[J].外语教学理论与实践,2015(1):66-75.
- [3] 许并生,宋大琦.20世纪《水浒传》思想研究及《水浒传》思想论析[J].东南大学学报(哲学社会科学版),2008,10(1):111-119.
- [4] 许敏琳,陈忠暖.基于行为计量学的《水浒传》36天罡星性格特征新解[J].华南师范大学学报(社会科学版),2008(1):67-72.
- [5] 陈劲松.走出费孝通悖论——谈社会学的方法之争[J].清华社会学评论,2017(7):161-176.
- [6] 彼得·赫斯特洛姆.解析社会:分析社会学原理[M].陈云松,范晓光,朱彦,译.南京:南京大学出版社,2010:14-26.
- [7] 赵京胜,张丽,朱巧明,等.中文文学作品中的社会网络抽取与分析[J].中文信息学报,2017,31(2):99-106.
- [8] 埃里克·D·克拉泽克,加博尔·乔尔迪.网络数据的统计分析:R语言实践[M].李杨,译.西安:西安交通大学出版社,2016:49-52.
- [9] Nahapiet J, Ghoshal S. Social capital, intellectual capital, and the organizational advantage[J]. Academy of Management Review, 1998,23(2):242-266.
- [10] 宋志红,武天兰,史玉英.明星作者网络位置与科研产出的关系研究——以“竞争情报”领域为例[J].情报理论与实践,2015,38(5):55-59.
- [11] 张鹏程,李铭泽,刘文兴,等.科研合作与团队知识创造:一个网络交互模型[J].科研管理,2016,37(5):51-59.
- [12] 王颖.当前接近中心性对关键研发者创造力的影响[J].情报杂志,2016,35(12):169-174.
- [13] 张晓黎,覃正.知识与合作网络结构洞对技术创新绩效的影响[J].软科学,2013,27(12):58-62.
- [14] 伍德里奇.计量经济学导论[M].费剑平,译校.北京:中国人民大学出版社,2015:177-178.
- [15] 陈晓辉.大数据时代的文学研究方法——基于弗兰克·莫莱蒂文学定量分析法的考察[J].文艺理论研究,2016(2):70-77.
- [16] Jenine K. Harris. An introduction to exponential random graph modeling[M]. SAGE Publications, Inc,2013:42-48.

- [17] M Morris,MS Handcock,DR Hunter. Specification of exponential-family random graph models: Terms and computational aspects [J]. Journal of Statistical Software, 2008,24(4):1-24.
- [18] 周雪光. 组织社会学十讲 [M]. 北京: 社会科学文献出版社,2003:114.
- [19] 迪安·鲁谢尔, 约翰·科斯基宁, 加里·罗宾斯. 社会网络指数随机图模型 [M]. 杜海峰等, 译. 北京: 社会科学文献出版社,2016:203-216.
- [20] 迪安·鲁谢尔, 约翰·科斯基宁, 加里·罗宾斯. 社会网络指数随机图模型 [M]. 杜海峰等, 译. 北京: 社会科学文献出版社,2016:9,12.
- [21] Daraganova G., Pattison P. E., Koskinen J. H., et al. Networks and geography: modelling community network structures as the outcome of both spatial and network processes [J]. Social Networks,2012,34(1):6 - 17.
- [22] 苏晓萍, 宋玉蓉. 利用邻域“结构洞”寻找社会网络中最具影响力节点? [J]. 物理学报, 2015, 64(2):1-11.
- [23] Douglas A, Luke. A user's guide to network analysis in R [M]. Springer,2016: 93-94.
- [24] 罗家德. 社会网分析讲义 (第二版) [M]. 北京: 社会科学文献出版社,2010:3.

Statistical Network Analysis of Water Margin Heroes

NIU Liang

(College of Economics & Management, China Jiliang University, Hangzhou 310018, China)

Abstract: [**Purpose/significance**] In order to explain the ranking of Water Margin Heros and the formation of interpersonal relationships. [**Method/process**] This paper analyzes the effect of water margin heroes network centrality on seat ranking by means of a multiple linear regression model and analyzes the causes of water margin Heroes interpersonal relationships through an exponential random graph model. [**Result/conclusion**] The results show that there is a positive influence on relationship between degree centrality and meso-centrality on ranking in the ranking analysis; geographical proximity, bureaucratic class, poor people and wealthy businessmen have a significant influence on network formation in the relationship formation analysis, and there is a positive influence on relationship between degree centrality and proximity centrality on the formation of character relationships in network structure.

Keywords: Network of water margin heroes; Spelman hierarchy correlation; Linear regression model; Exponential random graph model

(本文责编: 王秀玲)