

复旦大学社会发展与公共政策学院
2019~2020 学年第二学期期末考试试卷

A 卷 B 卷 C 卷

课程名称: 互联网与社会研究 ☐ ☐ ☐ 课程代码: SOCI130195.01

开课院系: 社会发展与公共政策学院 考试形式: 开卷/闭卷/课程论文/其他

姓名: 孙李媛 学号: 16307090164 专业: 社会学

提示: 请同学们秉持诚实守信宗旨, 谨守考试纪律, 摒弃考试作弊。学生如有违反学校考试纪律的行为, 学校将按《复旦大学学生纪律处分条例》规定予以严肃处理。

题号	1	2	3	4	5	6	7	8	总分
得分									

(以下为试卷正文或课程论文题目)

使用话题模型和社会网络分析分析青年群体微博数据集, 呈现青年群体的在线表达与在线互动结构的特点, 并在此基础上提炼青年群体的特征, 讨论发现的社会学含义。要求:

- 1、结合平时作业的发现, 提炼出较为聚焦的研究问题, 围绕清晰的研究问题展开分析和论述。
- 2、正确使用社会网络分析和话题模型分析技术。
- 3、能够对社会网络分析和话题模型分析结果进行一定程度上的整合。
- 4、能够对相关的文献做初步的回顾(如青年群体的研究、网络社会的研究、微博研究等文献), 研究发现能够回应现有文献。
- 5、表述清晰、规范, 文章结构完整、层次合理。
- 6、字数不少于 5000 字。
- 7、论文排版符合学校要求, 可参照附件模版

微博空间中的青年网络文化群体生态

一、研究背景及文献回顾

群体是社会学研究领域的重要议题。随着互联网技术的兴起,人们能够便捷地通过网络相互连接,社会互动、文化生产从现实空间拓展到了网络空间,群体研究也从现实社会拓展到了网络社会。社会学的群体研究主要有两类取向:一类关注作为概念和事实的群体所具有的基本性质和整体特征,通过身份认同、群体内外互动、文化生产等角度来探讨群体的定义、形成和存在,研究议题如特定群体符号边界的形成、特定群体亚文化的生产和实践,等等(方文 2005;周怡 2018)。另一类则运用群体的分析范畴,譬如进行内群体与外群体、初级群体与次级群体、正式群体与非正式群体等等划分,对现有社会结构、互动行为进行分类和理解(戚攻 2001)。

针对网络群体的现有研究主要采用前文所述的第一类取向,将网络群体作为一个整体关注其特点、互动和文化。数字化和虚拟化的互动和结构使网络社会具有自主、开放、多元的特点(孙伟平,贾旭东 1998),但也使网络空间中的群体带有认同感和归属感弱、互动持续性差,并缺乏共同身份和共同目标等特征。这些特征使网络群体的界定具有模糊和不确定性,因而其研究价值更多地被置于与现实群体的对比,揭示网络互动对现实群体的影响、重组和再造等等(戚攻 2001、2003)。网络群体的互动被认为是一种“机械团结”,这是由于网络群体大多因为兴趣爱好的相似性聚集,互动往往只停留在精神和信息层面,因而是浅层的、对前现代社会的回归(童星,罗军 2001)。尽管网络群体的互动中包含有机的分工合作的因素,但往往需要现实群体为基础或作用于现实语境,如现实中的业主群体通过在网络论坛中的互动促成维权事件(黄荣贵,桂勇 2009)。由此可见,受制于研究视角的局限性和网络空间自身的特征,当前研究所呈现的网络群体较为浅层、松散,多依附于现实社会,在对网络群体进行主体性、独立性和深入分析方面仍有较大空间。

不过,也正因为网络空间信息化特性,创造和生长于兹的网络文化相对而言更加繁荣,通过用词用语、互动关系成为网络群体的重要表征。许多研究聚焦青年网民群体,关注到如“丧文化”、“晒文化”、“日常迷信”等青年网络亚文化现象,并据此对青年网民群体的主体特性和网络空间整体的娱乐性、消费性加以理解(闫方洁 2015;王昱娟 2020;刘汉波 2020)。亚文化是相对于主流文化而言的概念,传统的亚文化研究主要关注亚文化与主流文化之间抵抗和收编的二元对立,现代以来逐渐更加重视由小众群体主体性带来的多元和差异(周怡 2018)。已有研究和经验观察已经证实,网络空间中产生和生长了多种多样的亚文化,但却常常以青年文化、网络文化一言蔽之,而较少关注小众群体的主体性,看到青年网络群体内部的分异。因此,本文将从文化的角度切入,围绕网络群体的主体性,通过网络空间中的文化分异,尝试展现青年网络群体并非铁板一块、而具有多元化和差异化,从而对构成这一群体的各个分群体形成更有主体性、更全面的理解,展现网络空间中的群体生态。

二、研究问题和方法

本研究的核心问题是:青年网络群体是否以及如何具有内部的多元和差异性?对应问题的假设是:网络空间中存在多个青年社群,主要体现为不同社群所关注话题的差异。

本文选择新浪微博平台上的博文和转发关系数据进行处理。新浪微博是较为常用的、具有极强公共性和社交性的社交媒体平台,也是当代中国青年网络空间表达和互动的重要渠道,目前已经拥有约5亿活跃用户。微博的产品设计允许所有微博用户在微博平台上几乎无门槛、平等、公开地发布原创图文、视频内容,以及对他人的公开内容进行转发。因此,微博上的

博文能够体现当代中国青年群体在网络空间中关注的话题、表达的态度等等，而转发关系则反映了青年微博用户对互动对象的选择以及互动行为特征。

本文采用大数据的研究方法，利用社会网络分析和话题模型分析技术处理上述微博数据。大数据的方法在文化社会学研究领域被越来越广泛地承认和采纳，一个原因是，由于文本分析方法的局限，既往文化社会学的实证研究往往只能在微观文化要素的层面，而中观和宏观层面的诸如文化生态或功能理论则难以获得足够的数据支撑。而大数据的方法却能够高效地获取一个网络空间话语场域中的所有或海量内容并进行分析，极大地解决了前述问题（Bail 2014）。

笔者所使用的数据库主要涵盖了 2011-2018 年间 100 位微博用户原创、转发的博文内容和转发博文的来源。由于社群基于相互间的互动关系，笔者首先筛选出数据库中存在历史转发行为的用户和博文，并以此为基础构建一个由被转发者指向转发者的社会网络，并进行进一步的社群分析。

由于社群分析仅关注互动关系，为了凸显网络中重要的互动关系，并获得更加清晰的可视化效果，笔者过滤了该社会网络中中心性较小，亦即地位低、权力小的节点和孤立节点，并计算网络的指标。在社会网络分析中，比较关键的指标是节点的中心性（用 degree 来衡量），中心性大小与节点的权力紧密相关，入度（indegree）和出度（outdegree）的大小分别代表了该节点信息获取的权力和信息发布的权力（平亮，宗利永 2010）。结合本文的研究问题，该社会网络中的入度（indegree）可被理解为节点在社群内部和外部的活跃度，出度（outdegree）可被理解为在社群内部和在外部社群的影响力，这两个指标均体现了该社群的互动特征。

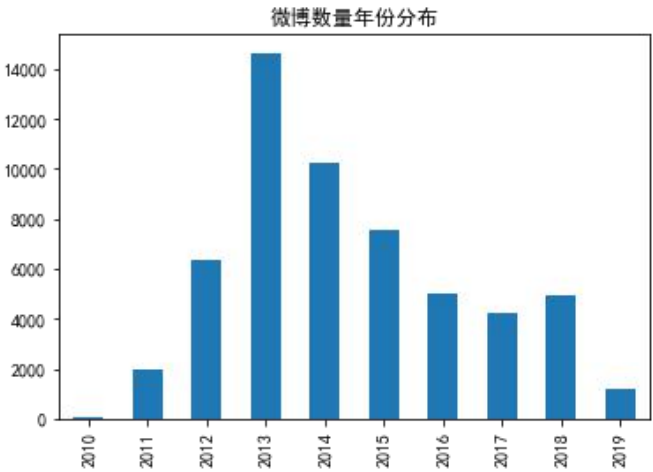
然后，笔者按照上述分析所得出的社群互动指标，将社群分别合并为不同类型的青年网络群体。然后，利用话题模型计算各个群体所关注的话题及其关键词，并结合对博文的质性分析给每个群体关注的话题命名和概括，呈现出每个群体整体上横向的文化框架。此外，笔者还筛选出群体内影响力较大（也就是出度最高）的核心成员，对核心成员的特征和博文做质性分析。

三、 数据处理结果及分析

1. 社会网络分析

（1）数据预处理

笔者从数据库中筛选出 97 位存在转发行为的用户和 54891 条博文作为分析对象。按年份对博文数进行统计，所得到的分布如下图：



从微博数量的年份分布图可以看出，2011 年及之前的微博数量较少，2010 至 2013 年间

微博数量快速增加,但在 2013 年后至 2018 年微博数量逐年减少。由于数据收集节点的影响,2010、2019 年微博数量统计并非全年数据,因此在本文的分析中不考虑在内。

(2) 社群侦测

将每位转发用户或原创用户作为网络的节点,将从原创用户到转发用户的关系作为网络的链接,并对网络中孤立节点和不重要节点的剔除,最终得出了一个有 1345 个节点和 2724 条链接的社会网络。对该社会网络进行社群侦测,发现了 18 个社群,其中 13 个位规模在 50 以上的大社群,5 个为规模在 50 以下的小社群。社群的划分主要依据节点之间的互动关系,联系紧密的节点被划分为一个社群,不同社群的节点互动关系相对松散。从可视化结果可以初步看出,不同社群的规模、互动强度、与其他社群的关系均有不同。为了使对社群的结构化分析更加具有文化意义,笔者选择了每个社群中出度(outdegree)也就是影响力最大的三位核心成员,对其发布的博文进行质性分析,以此概括每个社群的特征。笔者从影响力成员的性质、议题两方面对社群特征进行概括。前者分为官方(存在对应的现实组织、个人主体并被认证)、非官方(纯网络空间主体),以标明社群的正式属性强弱;后者为成涉猎的主要领域,如搞笑、资讯、情感等,以标明社群的文化取向。结果如下表:

编号	规模	核心成员	影响力成员性质
0	19	@高清电影院 @揭露地球 @凤凰卫视	非官方
1	208	@江宁公安在线 @夏影 @陈坤	官方
2	192	@新周刊 @顾扯淡 @我的极品评论	非官方
3	347	@DIY 设计我的家 @美剧贩 @林俊杰	非官方
4	9	@谢娜 @快乐大本营 @美丽人大	官方
5	212	@神店通缉令 @倒吸一口凉气君 @大神说	非官方
6	127	@微天下 @韩国 me2day @守望好莱坞	官方
7	357	@妖妖小精 @Houson 猴姆 @追风少年刘全有	非官方
8	217	@我的失眠读物 @小羊爱吃酱肘子 @草图君	非官方

9	168	@留几手 @新浪电影 @张嘉佳	官方
10	776	@全球热门搜罗 @全球热门排行榜 @这个绝对有意思	非官方
11	598	@人民日报 @央视新闻 @当时我就震惊了	官方
12	28	@张小娴 @历史最最最震惊 @詹姆斯吧官方微博	官方
13	14	@当红小花 i @复古写真馆 @公元 1874	非官方
14	81	@银教授 @知乎 @日本流行每日速报	非官方
15	37	@HeyWarWars @五行属二 @林宥嘉 YogaLin	非官方
16	253	@长安街知事 @一个抹布向前 @趣味贩卖机	非官方
17	426	@谣言粉碎机 @裸橙 @纸巾老师	非官方

社群 0 是一个规模只有 19 的小型社群，核心成员@@高清电影院主要分享电影资源，@揭露地球主要关注较为严肃、负向的时事热点新闻并发表评论，@凤凰卫视是电视台官方账号，发布时事热点新闻。三位核心成员的共性特征较少，难以捕捉社群特点，笔者将其概括为“关注时事新闻的电影爱好者社群”。

社群 1 是一个规模为 208 的中型社群。尽管核心成员为官方账号，但并不正式。@江宁公安在线背后的主体为@江宁婆婆，是一个在现实生活中就职国家机关、在微博平台上发布接地气、搞笑内容和评论的网络红人，而@夏影、@陈坤一个是漫画作家，一个是喜爱二次元、接地气的明星。因此，社群 1 可以被概括为“关注接地气官方账号的二次元爱好者社群”。

社群 2 也是一个规模 192 的中型社群。社群的核心成员@新周刊、@顾扯淡主要发布相对轻松的社会新闻、热点时事及评论，@我的极品评论则主要拾取微博评论区的精彩评论做集锦，三个核心账号尽管性质不尽相同，但都以相对轻松、娱乐性的观点输出和整合为主要博文内容，因此社群 2 可以被概括为“关注轻松型时事热点的观点接收者社群”。

社群 3 是一个规模 347 的大型社群。核心成员主要为发布家居设计灵感、欧美影视资讯等生活、娱乐内容，核心成员@林俊杰为知名台湾男性歌手，出道较早，粉丝年龄偏大，主要在微博发布生活日常。该社群核心成员的共性特征并不明显，反映出该社群的关注领域较为松散、多元，笔者将其概括为“关注家居、娱乐和明星的生活型准中年人社群”。

社群 4 是一个小型社群，偏好也相对明显，可以被概括为“关注综艺娱乐的人大学生社

群”。

社群 5 是一个规模 212 的中型社群，核心成员均为发布热门搞笑、奇闻趣事的非官方账号，同时会进行一定的评论等观点输出。可以被概括为“搞笑、新奇内容爱好者社群”。

社群 6 是一个规模 127 的中型社群，以发布国际资讯为主，内容涵盖社会新闻和娱乐明星。核心成员有两位是新浪平台的官方账号，分别发布全球热点新闻和欧美影视资讯，@韩国 me2day 则主要发布韩国娱乐资讯。这一社群可以被概括为“关注国际新闻、娱乐资讯的国际视野社群”。

社群 7 是一个规模 357 的大型社群，核心成员以非官方账号为主，发布的内容为娱乐资讯和搞笑段子，但与其它娱乐搞笑博主不同的是，@Houson 猴姆、@追风少年刘全有等账号发布了许多原创、独家内容，@妖妖小精也具备漫画、综艺翻译的原创内容输出能力，都在微博空间中具有一定的地位和话语权。因此，笔者将这一社群概括为“一手、独家搞笑内容、娱乐资讯的追随者社群”。

社群 8 是一个规模为 217 的中型社群，核心成员均为非官方账号，发布的内容较为生活、随意，主题模糊，涵盖了搞笑内容、情感文字、生活分享、时事观点、福利抽奖等等领域。较为突出的共性特点是，这些账号发布的内容明显地主要针对年轻女性受众。因此，这一社群可以被概括为“关注生活化内容的青年女性社群”。

社群 9 同样是一个规模为 168 的中型社群，核心成员多为官方账号，@留几手主要发布时事评论，@新浪电影发布电影资讯，@张嘉佳为男性作家，主要发布观点文字，整体社群氛围相对文艺、严肃。因此，笔者将这一社群概括为“思想者和文艺青年社群”。

与社群 5 相比，社群 10 同为以发布热门搞笑、奇闻趣事的非官方账号为核心成员的社群，规模却达到了 776，在所有社群中规模最大。社群 10 的核心成员具有非常明显的特点，即博文内容大多为搬运、转发，缺乏原创、观点等人格化特质，营销和商业色彩较浓，但信息整合能力强大、内容丰富多彩。笔者将这一社群概括为“全面、多彩搞笑内容的阅读者社群”。

社群 11 也是一个超大型的社群，核心成员特征明显，@央视新闻、@人民日报均为政府机关官方媒体账号，发布实时、权威、重大的全国性新闻，@当时我就震惊了则是微博平台上一个知名度较高、活跃较久的内容搬运博主，主要发布新奇、有趣、搞笑的内容。由此可知，该社群是一个“追求权威、严肃信源也兼顾娱乐放松的国家大事关注者社群”。

社群 12 是一个小规模社群，核心成员多为官方账号，@张小娴为香港女性作家，主要在微博发布情感文字，@历史最震惊主要发布奇闻趣事，@詹姆斯吧官方微博主要发布篮球运动员詹姆斯的资讯。核心成员的共性特征并不显著，笔者概括为“篮球、文艺和新奇内容关注者社群”。

社群 13 同样是一个规模仅有 14 的小社群，但核心成员均为非官方账号，@当红小花 i、@复古写真馆主要发布明星，特别是美女的图片和资讯，@公元 1874 则主要进行港乐、影视评论等观点输出，因此笔者将其概括为“爱看美女的明星娱乐爱好者社群”。

社群 14 还是一个规模 81 的小社群，核心成员多为官方账号，@银教授为认证编剧，主要发布原创搞笑段子，@知乎为问答社区软件的官方账号，主要分享软件上的时事热门问答，该软件定位群体偏向精英和知识阶层，@日本流行每日速报主要搬运日本轻松可爱的趣味资讯。笔者将这一社群概括为“关注轻松、可爱、娱乐，特别是日本相关内容的知识阶层成员”。

社群 15 依然是一个规模 37 的小型社群，非官方核心成员@HeyWarWars 主要发布影音、时尚类资讯，@五行属二主要转发搞笑内容，官方账号@林宥嘉 YogaLin 为男性台湾歌手，主要在微博上分享日常。这一社群与社群 13 相比，性别偏好更不明显，态度和观点输出更少，在明星娱乐之外也会关注搞笑内容，因此笔者将这一社群概括为“影音娱乐领域的路人关注者社群”。

社群 16 是一个规模 253 的中型社群，核心成员账号均为非官方，@长安街知事主要发布北京地方新闻和一些重大国家新闻，显示出社群中北京和北京人的影响力大。@一个抹布向前发布原创段子，@趣味贩卖机发布时尚资讯，笔者将这一社群合理概括为“关注时尚和搞笑的北京用户社群”。

社群 17 是一个规模 426 的大型社群，核心成员@谣言粉碎机为科普网站果壳网下的辟谣账号，主要进行生活中常见谣言的辟谣和科普，@裸橙为个人账号，经常发布自拍等生活内容，是一个容貌姣好、穿着打扮时尚前沿、偏二次元的博主，@纸巾老师则是暴走漫画的认证主编，主要在微博分享生活。可以发现，这一社群与社群 1 同样关注二次元，但不同点在于本社群氛围更加生活化、个人化，因此笔者将这一社群概括为“生活化的二次元同好社群”。

（3）阶段性结论：

笔者通过定量和定性的方式，对该微博转发用户社会网络中的 18 个社群进行了概括和分析。社群的划分由用户间的互动关系决定，社会网络中存在 18 个社群，从一定程度上说明网络空间中的青年群体并非铁板一块，不仅不一定相互连通，还形成多种、多元的社群。对社群核心成员进一步做质性分析，笔者发现了不同社群调性存在一定的共性，如 18 个社群中，有 5 个社群都以搞笑博主为核心成员，5 个社群都以资讯博主为核心成员，7 个社群以影音娱乐博主为核心成员。但是，更加细致的质性分析却能够发现，即便都是关注搞笑、资讯或娱乐的社群，不同社群在所关注的地域范围、信源的官方和权威性、内容的聚合和发布形式上都会有明显的不同，由此显现出微博空间中看起来相似的青年社群，实际也拥有不同的价值取向。由此，假设初步得以证明。

2. 话题模型分析

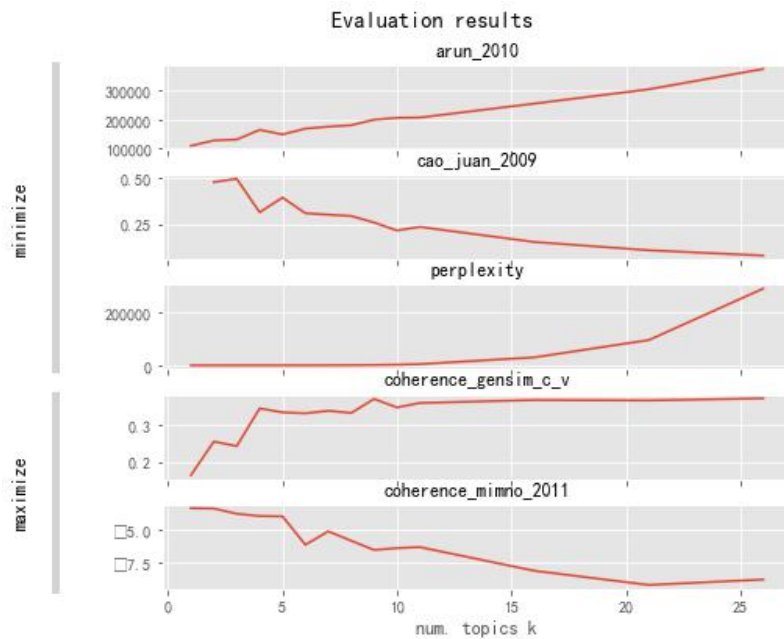
（1）数据预处理

为了进一步增加对社会网络结构的文理解，笔者对利用话题模型提取出社会网络成员所关注的话题和关键词，试图发现网络整体和每个社群的文化取向。

笔者首先在百度停用词表的基础上添加了主要由微博系统自动生成的词汇作为停用词，如“转发”、“微博”、“网页”、“链接”、“分享”、“图片”、“视频”等。在初步的词频统计中，出现了不同数量的“哈哈”等情感表达类高频词语，为了更好地突出话题性词语，笔者也对这类词语进行了过滤。

（2）话题模型

笔者对语料库进行 1-26 个话题进行模型拟合，并将不同话题数模型的一致性和混乱度指标进行可视化，得出结果如下：



从上图可以看出，话题数在 5 或 9 左右时，模型拟合的指标较好，一致性和混乱度指标综合水平最佳。通过对关键词可解读性进行评估，最终话题标题与最重要的 15 个关键词如下：

话题 1：线上于线下的社会互动	话题 2：家国正能量	话题 3：东亚影音娱乐	话题 4：二次元	话题 5：人生感慨
喜欢	中国	网页	努力	一个
孩子	真的	谢谢	告诉	发现
关注	一个	日本	好看	世界
朋友	感觉	电影	生日	很多
音乐	生活	简直	一句	一种
这是	时间	来自	银魂	上海
工作	幸福	下载	滴滴	真的
地址	可爱	海报	记者	别人
抽奖	两个	查看	比赛	东西
免费	老师	作者	拥有	时间
小时	希望	全国	鼻屎	卧槽
中国	美国	系列	最新	人生
活动	有人	设计	现场	作品
选择	故事	表情	能力	男人
公司	国家	公布	肯定	西安

结合关键词和对博文的质性分析，笔者对以上五个话题进行了理解。话题 1 中，重要的关键词“孩子”、“朋友”、“工作”、“公司”、“中国”等，反映出讨论这一话题的本土化、日常化的社会角色语境。“关注”、“地址”、“抽奖”、“免费”、“活动”、“小时”等关键词显示出在话题 1 的讨论中包含对线上互动行为的参与。“喜欢”、“选择”体现出对这一话题的讨论持较为积极、主动的态度。综合来看，话题 1 可以被概括为“线上与线下的社会互动”。

话题 2 中，重要的关键词“中国”、“美国”、“国家”体现出一种国家主义和全球化的语境，“幸福”、“可爱”、“希望”等关键词则表达了这一话题讨论强烈的积极乐观的态度倾向。

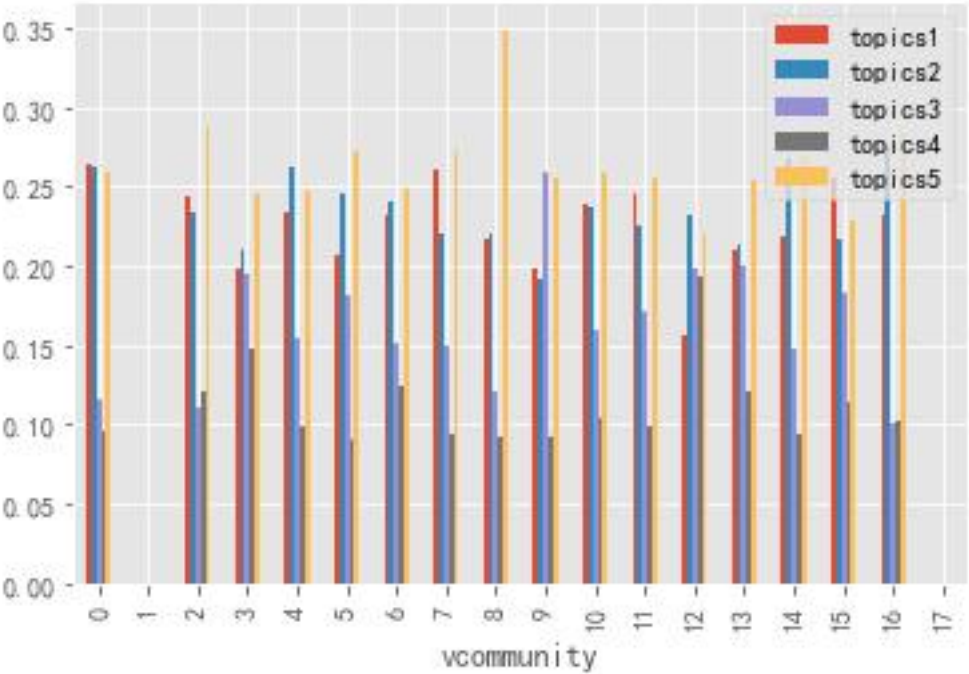
向。“生活”、“时间”、“故事”、“老师”等关键词显示出对这一话题的讨论是与日常生活、实践经验紧密相关的。因此，话题2可以被概括为“家国正能量”。

话题3的重要关键词“电影”、“海报”、“作者”、“系列”、“设计”展现出该话题的讨论对象为文艺创作，“全国”、“日本”等地理关键词说明讨论涉及的地域为本土和东亚地区，“网页”、“下载”、“公布”、“查看”则体现文艺创作的相关资讯和资源分享的语境。因此，话题3可以被概括为“东亚影音娱乐”。

话题4的关键词“银魂”、“好看”、“最新”说明话题中重要的讨论对象是二次元内容及评论，“鼻屎”（表情）、“生日”等关键词的出现提示话题的讨论者年龄相对较小，“努力”、“能力”、“肯定”同样也体现了积极的态度。话题4可以被概括为“二次元”。

话题5中，“人生”、“发现”、“时间”、“男人”、“世界”等关键词的出现说明话题5讨论的对象主要是人生感悟和日常感慨，同时出现了“上海”、“西安”地理关键词说明这一话题中上海和西安的用户占有较高的重要性，综合来看，话题5可以被概括为“人生感慨”。

每个社群的话题概率分布可视化结果如下：



话题1“线上与线下的社会互动”在社群中的分布相对平均，处于较高的水平。社群0、7、15在话题1上分布的概率高于其他社群，而社群12在话题1上的分布则较低。结合上文可知，更多倾向于讨论线上与线下的社会互动话题的社群是“关注时事新闻的电影爱好者社群”、“一手、独家搞笑内容、娱乐资讯的追随者社群”和“影音娱乐领域的路人关注者社群”，而不倾向于讨论该话题的社群则是“篮球、文艺和新奇内容关注者社群”。

话题2“家国正能量”在社群中的分布水平也相对平均，与话题1持平。社群14、16在话题2的分布概率相对较高，而社群9则较低。结合上文可知，更多倾向于讨论家国正能量话题的社群是“关注轻松、可爱、娱乐，特别是日本相关内容的知识阶层成员”、“关注时尚和搞笑的北京用户社群”，而不倾向于讨论这一话题的则是“思想者和文艺青年社群”。

话题3“东亚影音娱乐”在社群中的分布水平则有较大差异，整体水平相较于话题1、2、5偏低。社群3、9、12、13在话题3上的分布概率较高，而社群16的分布概率相对较低。结合上文可知，更多倾向于讨论影音娱乐话题的社群是“关注家居、娱乐和明星的生活型准中年人社群”、“思想者和文艺青年社群”、“篮球、文艺和新奇内容关注者社群”和“爱看美

女的明星娱乐爱好者社群”，而不倾向于讨论该话题的社群则是“关注时尚和搞笑的北京用户社群”。

话题4“二次元”在社群中的分布同样十分不均，且整体分布不高。社群3、12在这一话题上的分布相对高，其他社群则持相似的低水平。结合上文可知，更多倾向于讨论二次元内容的是“关注家居、娱乐和明星的生活型准中年人社群”和“篮球、文艺和新奇内容关注者社群”，而其他社群都对这一话题的兴趣度较低。

话题5“人生感慨”在五个话题中处于最高的分布概率，且在每个社群中都有相似的高水平分布。其中，社群8对话题5的讨论最多，社群12的讨论最少。结合上文可知，更多倾向于讨论人生感慨内容的是“关注生活化内容的青年女性社群”，而不倾向于讨论的则是“篮球、文艺和新奇内容关注者社群”。

（3）阶段性结论

从从话题分布的角度来解读，能够发现不同社群对话题1、2、5等大众话题的讨论上并没有太大的差异，而对话题3、4等小众兴趣话题的讨论则差异明显。由此可以看出，微博空间中的青年群体从整体上具有两面性，一面是对外界、对自己的普遍关切和表达，另一面则是主要通过兴趣爱好来实现差异化的自我形象建构。话题模型分析的结论部分地证明了笔者的假设，不同社群的青年确实会具有不同的话题取向和信息偏好，从而展现出不同的社群文化特征；但是这一差异主要体现在青年群体在网络空间中通过兴趣爱好进行的身分建构实践，现实生活、大众话题在网络空间的复现则并不具有太大的差异。

四、 总结

通过社群侦测和话题模型的分析，笔者部分地证明了自己“网络空间中存在多个青年社群，主要体现为不同社群所关注话题的差异”的假设。一方面，微博空间中的青年社群确实十分多元。网络空间中，青年用户并非是铁板一块、相互连通，而是形成了许多分异的社群，不同的社群在规模、核心成员性质、信息价值取向等方面都存在十分巨大的差异。社群的规模小则几人，多则七百余人，反应出社群之间差异化的互动关系。社群的核心成员性质也具有官方或非官方账号、原创账号或聚合转发账号、本土化或国际化等等差异。而不同社群的对搞笑、资讯、情感、观点、娱乐等信息类型和价值的取向更是不同。而另一方面，通过对每个社群所讨论的话题分布进行分析，也能够发现在日常生活、社会互动、家国时事等大众话题上，不同社群的青年具有相似的参与度。当青年群体涉及到自我独特身份形象的建构，即二次元、影音娱乐等兴趣爱好话题上的参与度才呈现出差异。

综上所述，本研究呈现出了与既往研究不尽相同的发现，即网络空间中青年群体的内部多样性。但是，在数据筛选、模型处理方面的技术水平不足，也使得本研究的效度打了折扣，无法实现最满意的分析效果，今后笔者将进一步完善。

参考文献:

- 黄荣贵、桂勇, 2009, 《互联网与业主集体抗争:一项基于定性比较分析方法的研究》, 《社会学研究》第 24(05)期
- 黄荣贵, 2017, 《网络场域、文化认同与劳工关注社群:基于话题模型与社群侦测的大数据分析》, 《社会》第 2 期
- 方文, 2005, 《群体符号边界如何形成:以北京基督新教群体为例》, 《社会学研究》第 2 期
- 平亮、宗利永, 2010, 《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》, 《图书情报知识》第 06 期
- 戚攻, 2001, 《从社会学理论域考察网络社会群体》, 《探索》第 02 期:77-80
- 童星、罗军, 《网络社会:一种新的、现实的社会存在方式》, 《江苏社会科学》第 05 期
- 戚攻, 2003, 《网络社会的本质:一种数字化社会关系结构》, 《重庆大学学报(社会科学版)》第 01 期
- 孙伟平、贾旭东, 2003, 《关于“网络社会”的道德思考》, 《哲学研究》第 08 期
- 周怡, 2018, 《“大家在一起”:上海广场舞群体的“亚文化”实践——表意、拼贴与同构》, 《社会学研究》第 5 期

Bail, C.A. The cultural environment: measuring culture with big data. *Theor Soc* **43**, 465–482 (2014).