

微信群社会结构演化规律研究：基于文本挖掘的案例分析

(南开大学商学院信息资源管理系 天津 300071)

摘要：微信已经成为移动网络用户进行信息交流、知识共享和协同工作的重要工具，但我们对于微信群的发展和成长仍然缺乏了解。本文利用 9 个真实的线上学术活动微信群数据，以虚拟社群理论和社会认同理论为基础，采用统计分析、社交网络分析和文本挖掘等方法，从群内和群间两个方面对微信群的演化规律进行研究。群内演化分析主要包括微信群的演化阶段、成员类型、成员间关系以及成员地位的研究，群间演化主要是针对群间关系以及群组间的资源整合与分配进行探讨。

关键词：微信群演化规律 社会认同 成员类型

Evolving Rules for Wechat Groups: A Case Study of Online Academic Activities

(Department of Information Resources Management, Business School of Nankai University, Tianjin, 300071)

Abstract: Wechat has become an important tool for information exchange, knowledge sharing and collaborative work for mobile Internet users, but we still lack the understanding of the development and growth of the wechat groups. In this paper, we carry on the analysis based on nine wechat groups used for online academic activities. Under the guidance of virtual community theory and social identity theory, we take statistical analysis, social network analysis and text mining to reveal the evolvement of wechat groups from two aspects, internal and between. Analysis of the changes within the groups mainly includes the evolutionary stages, members' types, relations and members' statuses. Intragroup evolvement principally focus on relations between groups, integration and allocation of the resources.

Keywords: Wechat Group; Evolvement Rule; Social Identification; Member Type

1 引言

移动互联技术开启了社交媒体的新纪元，使得信息的流动与共享呈现出不同于传统物理社区的新特征，如开放、瞬间传达、裂变式传播，等等。其中，增长最为迅速的移动即时通信工具是腾讯微信，2012 年 2 月上线，到 2014 年底用户数量已达 5 亿人¹。微信平台不仅提供一对一的私密交流、一对多的公众号传播，而且还提供多对多的朋友圈与微信群传播。微信群主要通过群成员邀请好友参与的形式建立，每个成员都可以邀请其它成员加入，这不同于以往 QQ 群只能由群主或得到授权的管理员才有权力

¹英媒：微信用户突破 5 亿 腾讯移动互联网地盘扩大.环球网,2015 年 3 月 20

邀请新成员加入的模式，体现出更加开放、去中心化、小世界传播的特点，群内成员的关系也不再只是具有高度中心性和密度的强关系结构，而是迅速扩展成员间的弱关系，并使得群的规模在短时间内以尚未确知的规律迅速增长。

当前微信群的应用已经融入社会生活的各个方面，如利用微信进行协同工作、开展网络营销、参与兴趣讨论等等。据不完全统计，截止 2015 年 4 月仅培训销售类、生活感悟类微信群数量达到 18738 个，其中 500 人以上的群占全部微信群数量的 65%^[14]。这些实际应用提高了工作效率，促进了信息与知识的扩散，但是随着群数量的增多，利用微信群传播淫秽信息、造谣、诈骗、聚众吸毒、甚至策划暴恐事件等不法犯罪行为的案件也不断增加。微信群的传播究竟呈现出什么样的特点？其增长和裂变空间具有什么样的规律？其内部的社会结构是否具有不同于传统物理社区和传统在线社区的新特点？微信群的传播对于成员的现实行为具有什么样的影响？对于这些问题的回答将有助于我们更好地掌握微信群的发展动态，有效预测其信息活动结果，从而更好地进行管理。但是，由于微信群传播的相对封闭性，后台大数据的获取又涉及到商业机密和个人隐私而难以获得，对其进行大规模的计量统计分析，而而困难重重。当前对于微信的研究尚处于起步阶段，侧重于探索微信中的信息传播机制^[3-6]、利用微信开展信息服务^[7-9]、以及分析微信使用中的用户信息行为^[10-13]等，而对于微信群则缺乏足够的关注。

微信群以移动性、碎片化、私密性的传播方式和迅速扩展社会资本的优势，成为传统虚拟社群的重要补充，对于群体性社会认同、线下集群行为动员、市场营销目标定位、圈群社会结构的形成以及社会文化的演变具有重要影响，因此，探索微信群的演化规律具有重要的理论意义和研究价值。为了有效探索微信群的演化规律，本文拟通过采集真实的微信群聊数据，以虚拟社群理论和社会认同理论为基础，利用统计分析、社交网络分析、文本挖掘等方法，探索微信群内外的演化规律。

2 研究的理论基础

2.1 社区理论

社区(community)是人类的共同体，在一些文献中又被称之为社群，是指在自然发展过程中形成的居住相近、具有共同利益、彼此相互需要的一群人。传统社区最重要的构成要素是共同性和互动性^[1]。社会学关于社区的研究具有较长的历史，较为有影响力的理论包括人文区位理论、社会体系理论、社区冲突理论和社区行动理论等。Park等人从生态学角度将城市区位竞争划分为侵入(invasion)、竞争(competition)、延续(succession)、驻扎(accommodation)四个阶段^[ii]。Coleman认为冲突是引起社区变迁的重要因素，其根源主要包括：①经济争端；②政治争端；③价值观的冲突。

2.2 社会认同理论

社会认同理论(Social Identification Theory)^[16]反映了社会群体意识构建的渐进过程，是研究现实社群或虚拟社群的重要理论基础。社会认同理论认为个体通常都是基于他与他所认同的群体的相似程度以及该群体与其他群体之间的非相似性来进行自我归类，这种自我归类发生在个人特征与社群特征的比较过程中^[17]。在自我归类之后个体会将自己归于与其相似的群体中将自己视为其中的一员，由此可知，个体对该群体感知到的相似程度越高，他对该群体的归属感也就越强。根据 Algesheimer 等人的定义，社群身份是个人对所属虚拟社群的感知^[18]，在成员与社群不断交互的过程中，他的个人身份会随着个性化解体而不再凸显，

而情感认同逐渐将其转化为社群的依附感和归属感^[19]。在缺乏实物元素的网络空间中，社群身份对于塑造成员的积极参与行为具有重要作用。

认同同样能够促进个体对群内成员的帮助和支持行为^[20, 21]，这些支持行为包括推动社群的发展和主动付出更多的努力来获得社群的共同利益。同时，基于认同的关系能够促进个体成员的自我管理，归属感强的成员会关注影响社群的相关问题，带有强烈认同感的成员则会与相关受众积极交流并传达社群的正向信息^[22]。相比于传统虚拟社群要求成员主动参与的形成机制，微信群既可以允许用户通过群的二维码进行验证加入，也可以通过群主或群内成员邀请的形式构建。因此，微信群成员社群身份的构建及其社会认同感可能需要经过更长的时间的群内交流才能建立。

2.3 在线社区研究

在线社区(online)又被称为虚拟社区 (Virtual Community)，是“当足够多的人在网络上进行长时间的公共讨论，并通过人际情感在虚拟空间形成个人关系网络时所出现的“社交聚集集体”(social aggregations)”(Rheingold,1993,p5)。在线社区是一种新型的社会组织，具有表达自由、缺乏集中控制、多对多传播、成员自主互动等特点^[1]。传统的在线社区包括博客、论坛、微博、社交软件、维基、贴吧、大型企业内部交流社区、视频网站等。对于在线社区的研究涉及到心理学、社会学、企业管理、计算机科学、信息系统、情报学等多个领域，研究内容包括技术媒介与机制、物理与在线虚拟社区、用户生成内容的价值、社区成员间的关系与情感、在线社区的发展、实施、效果与应用等。

在线社区给个人带来的好处包括信息交换、社会支持、社会互动、时间与地点的零活性、持久性等 (Maloney-Krichmar & Preece,2005; Johnson & Ambrose, 2006; Preece, 1998,1999)。刘永丹采用基于 Agent 的建模和仿真方法考察了微信群内的舆情演化的特点，研究发现，微信群内的舆情信息存在初期快速扩散到中期稳定传播，再到后期逐步消亡的周期性特征，群的拓扑结构决定着舆情演化的初步效果，同时，群的活跃度也影响舆情的传播速度^[15]。Millen 等人 (2002) 认为在线社区为组织带来的益处包括消费者忠诚、职员沟通与信任、可见性与声誉、生产力的提高等。

Lazar Preece(1998)发现不同社区的四个维度，包括属性、支持软件、与物理社区之间的关系以及边界。其中属性包括目标与兴趣 (Kim,2000)，家庭与生活方式，工作，游玩，精神与健康，政治，商业交易，热情交互与情感联系，共享活动，共享资源，支持，惯例、语言与礼仪，规模，匿名程度与收入来源；与物理社区的关系包括基础 (通常是面对面的)，城市，政府，教育，在线学术社区，爱好，匿名与角色扮演，健康，犯罪等)。KoI 等人对在线社区的考察主要通过研究用户的使用粘度、参与动机以及影响信息分享的因素等，探索保持在线社区活力的方法和途径^[2]。

在线社区的发展具有阶段性，Wegner 等 (2002) 识别出了在线社区构建的五个阶段：潜在、集结、成熟、管理与转型。Malhotra 等 (1997) 建议了三个阶段：开始在线社区、鼓励早期交互、走向自我持续的交互环境。Iriberry 与 Leroy (2009) 将在线社区的生命周期划分为开始、产生、成长、成熟与死亡五个阶段，并分析了各个阶段的成功因素²。上述研究对于在线社区的发展过程进行了有效的刻画，但是却认为在线社区的开端基本上都是基于现实的需求，因而无法用来考察不同在线社区之间的关联。现实观察表明，相比于传统的在线社区，微信群的产生更加容易，常常由已有的微信群裂变或衍生产生，因此有必要分析微信群间的演化规律。

² Iriberry, a. & Leroy, G. A life-cycle perspective on online community success. ACM Computing Surveys, 2009, 41(2), Article11.

基于上述理论研究成果,本文拟重点回答以下问题:(1)微信群的成长经历了哪些阶段?(2)群内成员的类型、角色及相互间的社会关系如何?(3)群的裂变与衍生具有什么样的规律?(4)微信群与现实社会的关系如何?为了回答这些问题,本文采集了真实的线上学术活动微信群数据,以虚拟社群理论和社会认同理论为基础,采用统计分析、社会关系网络分析和文本挖掘等方法,从群内、群间以及微信群与现实活动的关系三个方面对微信群的演化规律进行分析。群内演化分析主要包括微信群的演化阶段、群内成员类型、群内成员间关系以及成员地位的研究,群间演化主要是针对群间关系以及群组间的资源整合与分配进行探讨,而群与现实的关系则主要考察群内讨论对于现实活动的影响。

3 数据来源与研究方法

本研究选取了研究者所加入或组建的 9 个微信群作为研究案例,采集了真实的群内交流数据作为文本分析的对象。这 9 个微信群建立的主要目的是讨论学术问题或筹办学术活动,因而既有线上讨论,也有线下活动,比较适合本文的研究目的。9 个群的群成员主要是来自全国多所高校和研究机构相关领域的研究人员及少量政府机构的工作人员,成立的时间跨度在 2014 年 4 月 28 日至 2015 年 4 月 3 日之间。其中,成员人数最多的群也是成立时间最早的群,群成员共有 200 人,而成员最少的群只有 23 人。为了遵守研究伦理,保护研究对象的隐私,本文对群和群内成员进行了匿名化处理。数据分析只涉及群的特征与属性变化,并不涉及具体的交谈内容。按照群成立的时间将群编号为 QA 至 QI,各群的基本信息如表 1 所示。

表 1 微信群基本信息

群名称	成立时间	成员数(人)	成立目的	总发言量(条)
QA	2014 年 4 月 28 日	200	线上学术交流	8241
QB	2014 年 5 月 4 日	85	筹办学术活动	7268
QC	2014 年 5 月 25 日	78	线上学术交流	1651
QD	2014 年 6 月 30 日	29	信息扩散	106
QE	2014 年 10 月 24 日	45	筹办学术活动	655
QF	2014 年 11 月 3 日	30	筹办学术活动	1131
QG	2014 年 12 月 9 日	72	筹办学术活动	1956
QH	2015 年 3 月 4 日	86	筹办学术活动	5061
QI	2015 年 3 月 31 日	23	讨论特定议题	1185

本文采取的研究方法主要有统计分析、社会关系网络分析与文本挖掘方法,分别用于统计群的属性随时变化规律,群成员的社会关系结构,以及群内交流的主题演化规律。

4.群规模随时间变化规律

群的规模和发言量随时变化的情况可以反映群的演化规律,有可能揭示群的生命周期。由于成员在加入群聊之后可以进一步邀请自己的好友加入,因此我们首先对群成员加入群的时间进行统计,构建群成员数增长的时间序列热图,如图 1 所示。从图 1 中可以看出,群的扩张主要发生在群建立的初始阶段,群成员人数从建立到达到完全饱和的 50%,即总体用户量一半,只需要约 3 天的时间,而群的扩张在后期的整个阶段中都会持续进行。其中,QA

和 QC 是以线上学术交流为目的建立，QB 及 QE 到 QH 是为筹办学术活动所成立，这几个群在成长的过程中一直都有新成员的加入，而 QD 是对少数成员传递通知，以传播和扩散信息为目的，QI 是以讨论特定议题为目的，群 QI 的建立初期就明确不再添加新的成员，因此，群人员的增长与群的成立目的紧密相关。

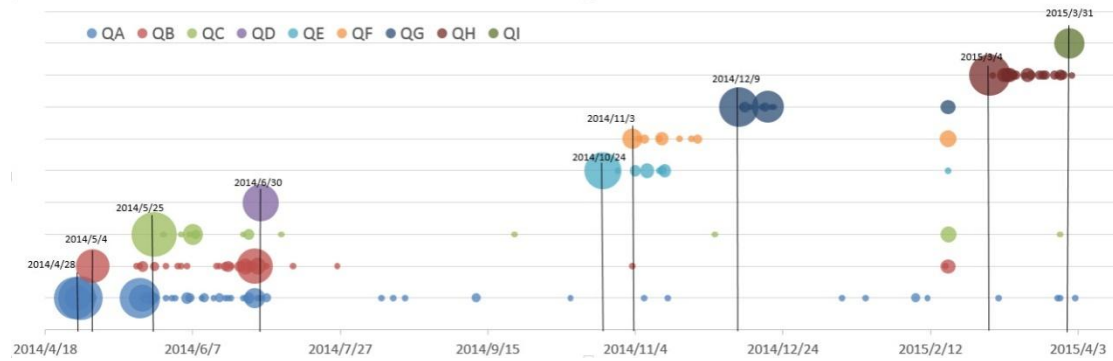


图 1 群成员数增长的时间序列热图

以天为单位对 9 个群的发言量进行统计，绘制微信群发言量的时间序列热图，从而展现各个群在整个时间跨度上的活跃程度，如图 2 所示，其中气泡的大小代表当天群内发言的数量。

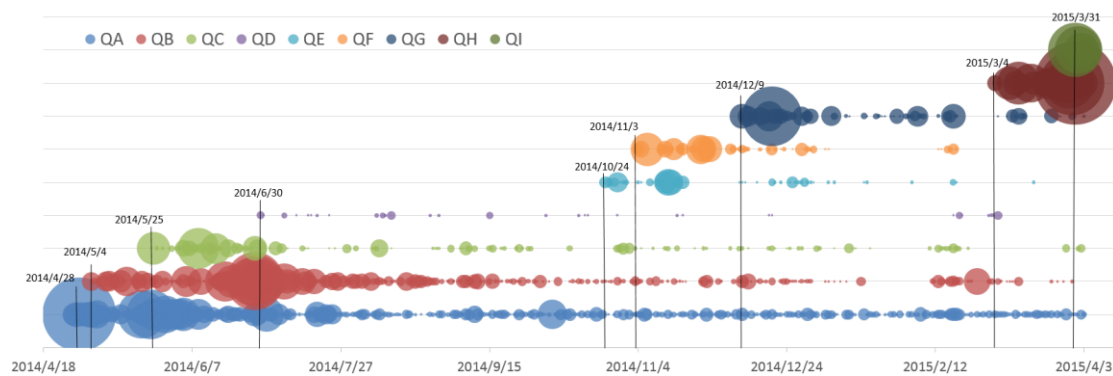


图 2 微信群发言量的时间序列热图

从图 2 中可以看出，在微信群建立的初期，用户发言讨论最多，随着时间的迁移逐渐消减。其中，以线上学术交流和筹办学术活动为目的的群持续活跃时间最长（QA-QC、QE-QH），而以信息扩散和讨论特定话题为目的的群活跃时间较短（QD、QI），特别是 QD 的发言量时间序列中并不存在明显的热点，发言时间也较为零散，因此，成立目的对于群的持续活跃时间存在直接的影响。

对 9 个微信群的发言量进行统计，并以小时为单位对各个时间段的发言量占总体发言量的比例进行计算，用户发言时间段的分布情况如图 3 所示。

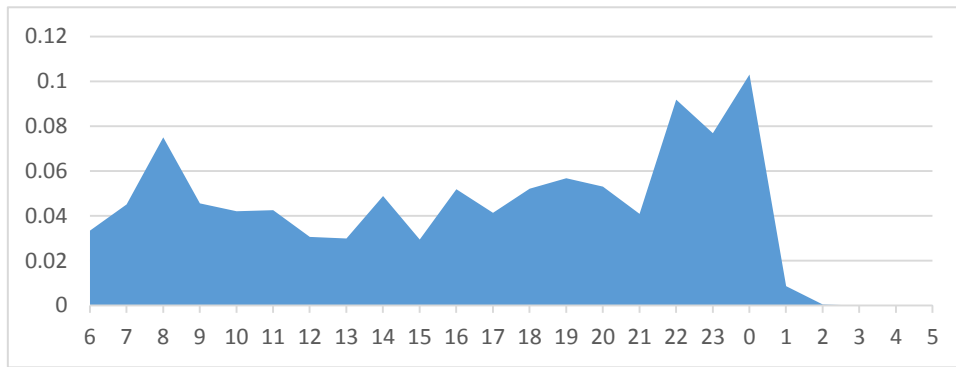


图 3 用户发言时间段分布情况

与微博不同，微信群的用户在一天中清醒的时间段内一直保持着活跃的互动。在 8 点、11 点、14 点、16 点、19 点以及夜间的 22 点、0 点显示出明显的互动高峰，12 点至 13 点之间及 15 点，用户互动频率下降，分析其原因在于工作疲惫，精神不佳。根据美国东北大学 Alan Mislove、Sune Lehmann 等人针对 twitter 数据所进行的情感分析结果显示，人的情绪会在下午 3 点到 4 点间跌入低谷，这也能够解释案例中用户互动频率下降的情况^[23]。2 点之后至 5 点，用户进行睡眠，互动量为 0。这是因为微信传递的方式和提醒方式不同于微博，微博在发言互动过程中较少因消息的到达声音影响到用户的正常生活，而微信发言则会直接提醒用户，影响用户休息。

4 群内演化分析

为了解微信群内的演化规律，我们选择群 A 的实际数据进行分析，选择群 QA 的原因在于，群 A 存在的时间跨度最长（约 1 年的时间），用户数和发言数据量最大（用户数为 200 人，发言数为 8241 条），发展最为成熟。

4.1 群内演化阶段分析

微信群网络在建立之后，内部成员处于一个封闭的全互联有向网络中，外部成员无法与网络内部沟通，内部成员可以选择对群内所有成员广播，也可以在群内与某一用户单独对话（通过<@用户名>的方式提醒某一用户注意某条特定信息），内部成员是平等互通的，但通过实际数据分析我们发现，用户在群中的地位 and 活跃程度都有所差异，以全联通有向网络建立的微信群网络结构不能有效描述微信用户在群中所处的位置和行为，同时，我们认为以<@用户名>的方式进行交流的两具有强关联关系，这种关系能够更好地反映微信群网络的结构特点。因此，我们对微信群网络的构建选择以用户为节点，以用户的发言量表示节点权重，用户被其他用户@的关系作为有向边，@的次数作为边的权重进行赋值。

通过对案例数据进行分析，我们将微信群的演化过程划分为 6 个不同阶段，包括发起、扩张、活跃、稳定、沉寂和分裂，并根据不同时间段内微信群 QA 的实际数据构建微信群网络结构演化图，如图 4 所示。

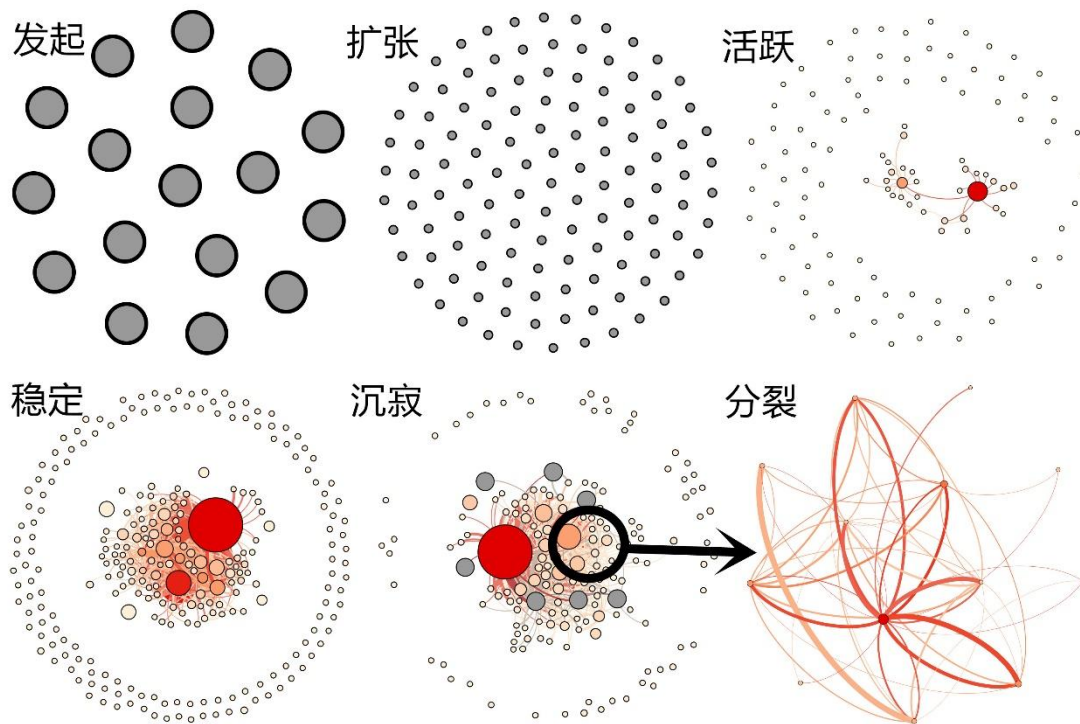


图 4 微信群 QA 网络结构演化图

在微信群的发起阶段内，群主通过邀请好友加入群聊构建初始的群网络，各成员之间处于平等关系，均匀分布在群网络中，互相之间可能互为好友或是陌生人。初始成员继续邀请新的好友加入，群规模逐渐扩大，进入群的扩张阶段。随发言量的逐渐增加，群内成员互动程度加深，逐渐形成相对稳定的中心子网。在议题结束之后的较长时间段内，群内用户的活跃度逐渐降低，并且部分用户选择离开，微信群进入沉寂阶段。在整个演化过程中，群中的活跃用户逐渐与其他成员建立强关联关系（如互相添加为好友），群的使命完结并分裂，分裂的形式可能包含网络内部成员加为好友转为私下交流或是衍生出新群。

4.2 群内成员类型分析

对群 A 中所有用户的发言量和被@数进行统计，并计算各自占总体的比例，按降序排列，结果如图 5 所示。



图 5 群 A 用户发言数与被@次数占总体比例情况

群 A 中发言人数共 157 人，占群内总数的 $157/200 \approx 78.5\%$ ，发言用户中前 20%（约 31 人）的发言量之和占总发言量的 $6731/8241 \approx 81.7\%$ ，明显符合长尾理论。从图 5 中可以看出，用户发言量所占总体比例与被@次数的曲线趋势相当一致，表现为用户发言次数越多，被其他用户指定交流的次数越多，越容易与人进行沟通，建立强关联关系。虚拟社区所提供

的互动环境和互动支持，促进了社区成员之间的交流活动，帮助成员形成自我以及获得和展示个人身份^[24]。一旦社区成员在互动中得到别人的角色支持，喻示着自我展示取得预期响应，从而建立对虚拟角色的自我认同；同时，研究者已经证明社区成员之间的互动允许会让成员（1）感知社区更加有吸引力^[25]，（2）而且产生对社区的归属感^[26]，这些增加的互动活动，促进成员把自己看成是该虚拟社区成员中的一份子^[27]。因此，虚拟社群中的互动活动构成了网民认知、身份建立和情感联结的基础。

毛波和尤雯雯通过对 BBS 虚拟社区相似版面讨论成员的数据进行聚类分析，将成员类型分为 5 种，领袖、呼应者、浏览者、共享者和学习者^[28]。通过对微信群稳定阶段的网络结构图进行分析，我们将成员划分为核心分子、活跃分子、半活跃分子、沉默围观者和疏离者几种类型，如图 6 所示。

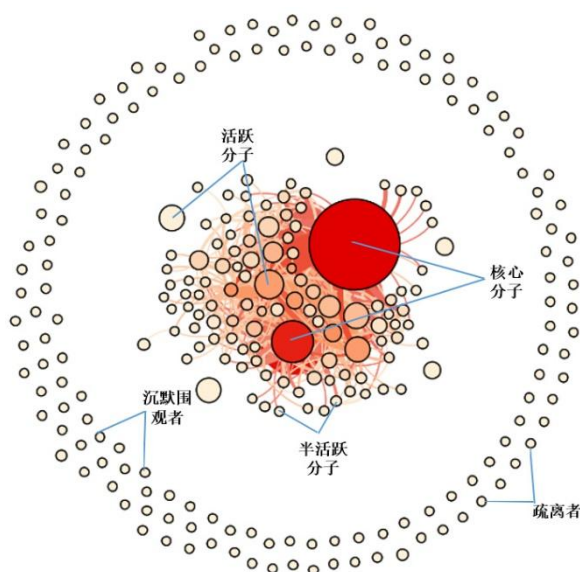


图 6 微信群成员类型示意图

核心分子和活跃分子组成了活跃区，其中核心分子一般较少，有一个或几个，这些成员的发言量和被@的次数甚至会达到总体的 10%以上。活跃成员也会经常发言或互动，围绕几个核心分子形成几圈“涟漪”状结构。活跃分子、半活跃分子相比较于核心分子其节点度较小，但仍处于网络的中心位置，活动范围较为集中，部分活跃分子和半活跃分子不存在与他人的直接交流关系，而是选择对群体成员广播，这取决于个人的信息行为习惯。沉默围观者虽然很少发言，但是会关注群的信息，必要时会参加讨论。沉默围观者受到中心话题和活跃用户的引力作用，聚集在核心分子和活跃分子的周围，这几圈“涟漪”常常相互交叉叠加，构成一个相对稳定的社会关系网络。而疏离者则在中心网络的四周围观，基本上游离在讨论之外，很少关心群里讨论的内容。

从沉默围观者和疏离者的入群时间来看，主要集中在群发展的中后期，与他们的发言时间进行对比发现，二者很难进入群的核心网络中，与其他成员的交流较少。分析其原因在于，首先，围观者和疏离者往往是由群内的半活跃成员和其他围观者邀请入群，并且是在群建立的过程中逐渐加入，除邀请者外，群内其他成员对他们而言都很陌生，很难与其建立联系，因此选择观望。此外，他们对群建立的背景和前述对话都不了解，因此无法融入群的情境之中。

4.3 群内成员关系分析

微信群的成长过程与其他类型的虚拟社群较为不同，群内成员都是由其他成员邀请进入，也就增加了一个初期过滤的过程，群成员在选择被邀请者的过程中就对被邀请者的个人信息与群信息进行了匹配，因此被邀请者与该群体已经具备了一定的相似性。因而，微信群成员在加入群的初始阶段，与其他群用户之间具有较高的相似度，个体能够迅速建立强烈的归属感，这也就解释了为何微信群在成立初期能够迅速的进入活跃阶段。我们对成员加入前后的相关 10 条记录进行文本分析，通过文本挖掘，人工筛选出能够反映成员加入目的的词语，结果如表 2 所示。

表 2 成员加入前后用户发言的关键词统计结果

关键词	词频	关键词	词频
平台	90	调查	30
学习	83	培训	21
项目	76	跨界	16
合作	74	分享	15
领域	58	课题	14
研讨	56	互通	12
共享	31	同仁	7

对维基百科开源社群中协同信息行为的研究发现，用户贡献知识的主要激励因素是声誉，而非帮助他人的愉悦感，社群内部大部分的成员都是务实的利他主义者^[29]。因此，只要虚拟社群提供条件以吸引务实的利他主义者贡献知识，就能保证社群的持续性^[30]。从分析结果中我们可以看出，微信群为学术交流提供了很好的互通交流平台，成员也希望通过这个平台结交更多的相同或相似领域的同仁，获取更多的合作机会，其主要是以研讨、调查、培训、寻找科研课题等为目的，同时，参与者也希望获取多领域跨行业的知识和交流机会。作为封闭的社群网络，微信群成员的特性与开放社群成员截然不同，他们之中大部分成员都是务实的利己主义者，成员间关系的建立更多的是一种资源互补的关系。

我们把虚拟社群提供的条件看作资源，成员之间的交流则是资源的互换整合，虽然半开放的微信群也经历着成员不断加入退出的过程，但这个过程相对于开放社群的成员变动而言非常微弱，因而微信群内的资源的总量和更新频率在扩张期过后基本保持不变，资源的交换也在活跃期中基本转换完成，因此，微信群在经过一段时间的稳定阶段后必将进入沉寂阶段。当新成员加入或者有人引入新话题时，群内又会产生短暂的新的活力，但当新话题与原主题群的话题关联性不大或是深入演化时，旧成员的知识结构不再能够满足新话题的关联需求，即不再有新知识的产生，这时，新话题的产生往往预示着新群的建立。

4.3 核心成员的地位

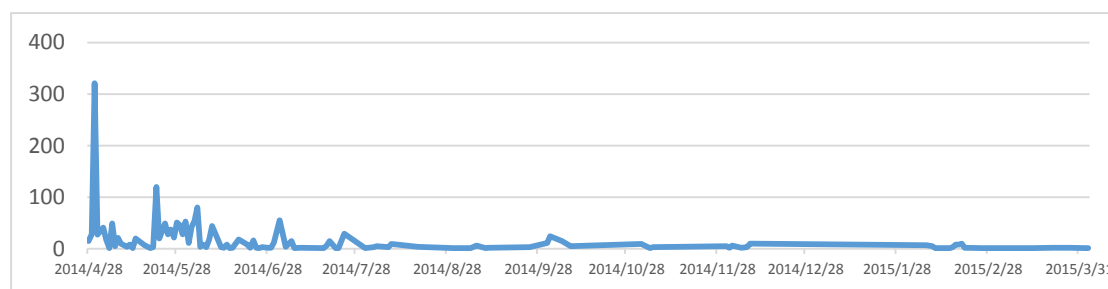


图 7 核心分子发言量统计图

对核心分子发言的时间进行统计,我们发现,微信群核心分子的地位是在群建立的初始阶段形成的,而随着群的扩张和逐渐消寂,核心分子不再活跃,而是选择在其他新群活动或者主动创建新群。分析可能原因在于,首先,以往的研究显示,人与人之间的信任程度对于虚拟社群成员的知识共享和知识收集行为具有强烈的积极影响,同时,信任对于资源交换和创新的情境也具有重要作用^[31]。而随着群的不断扩张,群成员不断增多,非本人邀请的成员越来越多,成员间的不信任因素逐渐加重,因此,某些核心成员和活跃成员选择缄默,不再继续发言讨论。其次,对于群初始建立的目的已经完成,对于某一学术问题的探讨不再有新的观点和问题产生,或者是某一线下活动结束,同样也就意味着群的使命结束。再次,核心成员间发生冲突,由于无法产生统一的观点,不能达成共识,情感分裂而选择缄默。最后,成员离开或加入新群,从而导致了话语对象的缺失。

对核心成员在网络中的地位进行研究,我们将沉默围观者、疏离者以及没有与其他成员联通的群成员剔除,按照用户发言量排序对用户进行编号,构建联通的有向群中心网络,其中节点大小为度值,边表示用户间相互@的次数关系,如图 8 所示。利用 Ucinet 计算各节点自我网络密度参数(EGO NETWORKSDensity Measures),部分成员的计算结果如表 3 所示。

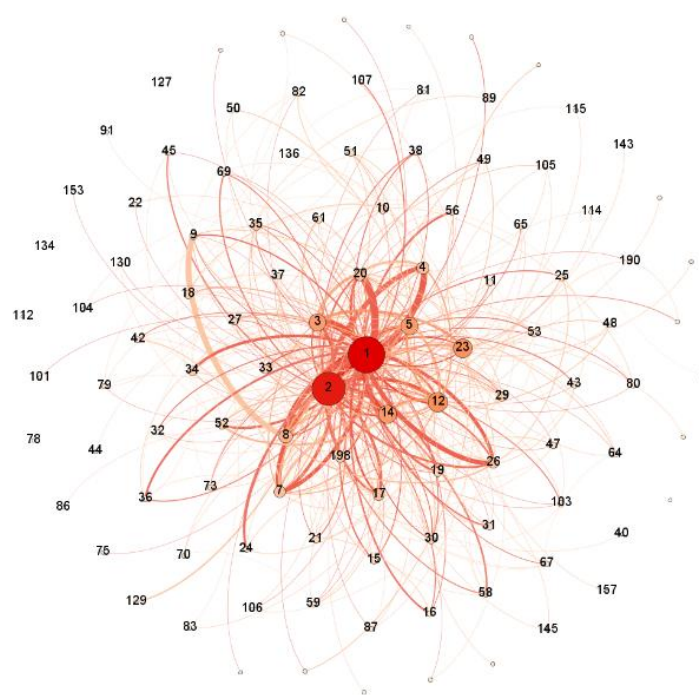


图 8 微信群中心网络结构图

表 3 微信群成员自我网络密度参数

成员编号	Size	Ties	Pairs	Densit	nWeakC	pWeakC	2StepR	ReachE
14	17	52	272	19.12	4	23.53	79.63	32.7
4	11	40	110	36.36	2	18.18	84.26	33.83
1	51	234.00	2550.00	9.18	7	13.73	97.22	18.921
2	29	148	812	18.23	2	6.9	91.67	22.6
7	13	58	156	37.18	1	7.69	93.52	31.27
8	13	33	156	21.15	3	23.08	74.07	37.38
12	23	116	506	22.92	2	8.7	85.19	23.77

结果显示,用户 1 在网络中形成了规模最大的个体网,与其直接沟通的成员 51 人,其个体网中成员关系数是所有其他成员个体网中最多的,也就是说,用户 1 在群内的交流中,主要是以扩散信息为主。二步网络值(2StepR)说明用户经过两步即可达到的用户范围,用

户 1 的计算结果为 97.22%，表明用户 1 在网络中扩散信息的能力最强。但用户 1 的个体网密度相对较小，只有 9.18%，也就说明其个体网中的成员间交流较少，主要以接受信息为主。因此，主体在微信群中与较多用户互动沟通能够扩大自身的话语空间，推动自身的信息传播和资源扩散，但相应地也在一定程度上丧失了与他人深入交流的机会，而成员较少的个体网中，如用户 4、7 所构建的个体网，成员在 10 人左右，其个体网密度都较高，因此可以证明其内部的互动更加频繁，交流更加深入，有助于问题的探讨，相关研究显示，微信群活跃度最佳人数是 11 到 40 人，这一规模最适合沟通，其使用度也最高^[32]。

社区认同理论认为，一旦社群成员将自我归类为某个社会群体，其群体成员的身份意识将会影响其群体内的偏私行为，其核心的解释是个体为了满足自尊的需求激发了其社会认同的动机，突出的是身份意义的价值层面。当网民知觉到自己归属于某一虚拟社区，将引发一个虚拟社区身份的类化过程，激活他/她的群体身份，形成该社群的集体意识^[33]。从分析结果来看，核心成员在微信群中的主要功能在于传播信息，激发群成员的群体身份，构建社群的集体意识。Bagozzi 和 Dholakia 的研究发现，只有社会认同和群体规模对于虚拟社群用户的参与行为具有重要影响^[34]，因此，核心成员在不断构建自我地位的同时，对于社群的启动和演化也起着非常重要的作用。

集聚系数(clustering coefficient) 是用来刻画网络节点的邻居节点间相互连接的稠密程度，反映了复杂网络内部集团化的程度。网络平均集聚系数是所有节点集聚系数的平均值，具体公式 1 如下：

$$C = \sum P(k) \times C(k) \text{ 公式 1}$$

式中， $C(k)$ 为局部集聚系数， $C(k) = |E_k| / (k(k-1)/2)$ ，其中 $|E_k|$ 为节点度为 k 的邻居节点中平均连接数。统计研究表明，大规模网络普遍具有明显的集聚特性，多数网络的平均集聚系数远大于 $O(N-1)$ ，部分网络高达 0.8 至 0.9，根据理论推测，对于很多网络来说，当网络规模 $n \rightarrow \infty$ 时， $C \rightarrow O(1)$ ^[35]。计算微信网络的平均集聚系数为 0.314，结果表明，网络的集团化程度较低，即用户的集聚较为稀疏。与前述结果结合来看，核心用户可能形成一个较小的子群，但在总体网络中的影响力较低，很难把持独立的话语权。

5 群间演化分析

学术微信群的成员往往来自具有相同或相似研究领域的学者，其内部成员的重叠程度可以反映出两个社群话题的相似程度，我们利用 Jaccard 系数计算不同微信群之间的成员相关度权重，对于任意两个微信群 A 和 B，公式 2 表示社群 A 和社群 B 的 Jaccard 系数计算公式，其中 $C(A \cap B)$ 表示群 A 和 B 共有的成员数， $C(A \cup B)$ 表示群 A 和 B 全部的不重复成员集合，根据微信群成员相关度矩阵建立群间关联网络，为更清楚显示各群之间的演化关系，通过实验观察，我们选择边权重大于 0.139 的路径显示，结果如图 9 所示。

$$JC(A, B) = \frac{C(A \cap B)}{C(A \cup B)} \text{ 公式 2}$$

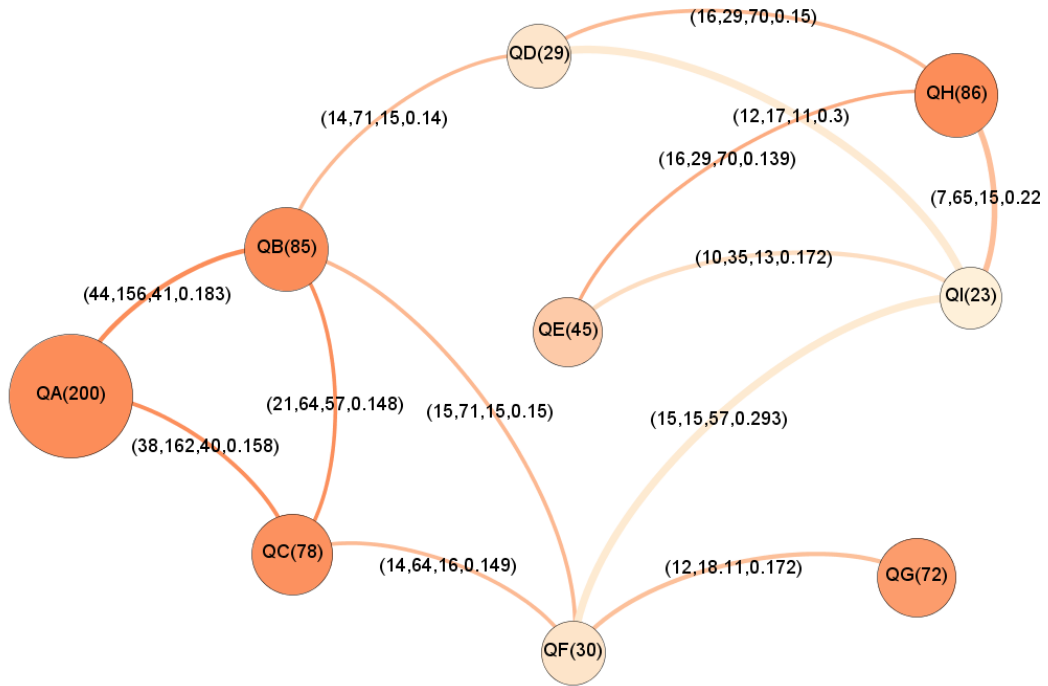


图 9 群间关联网

其中，各节点表示不同的社群，节点大小为群的成员人数，标签为群编号及成员人数，边表示群间成员的关联度网络节点间的边表示的是两个群所包含成员的相关关系，为更清晰显示群间关系，将边标签定义为 $(C(A \cap B), A - C(A \cap B), B - C(A \cap B), JC(A, B))$ 。其中 $C(A \cap B)$ 表示群 A 和 B 共有的成员数， $A - C(A \cap B)$ 表示群 A 未加入到新群 B 的人数， $B - C(A \cap B)$ 表示在群 A 的基础上群 B 新增加的人数， $JC(A, B)$ 为群 A 和 B 的 Jacard 系数。

根据本项案例中微信群成立的背景可知，各群都是围绕一个相似的研究主题在不同的时间阶段建立起来，从图 9 中我们发现，从 QA 到 QI 的过程中存在多条路径，经过的节点包括 QB、QC、QD、QE、QF、QH，但 QA 和 QI 之间没有直接的关系，表明两者之间的成员重叠程度很小，也就说明从旧群到新群，成员组成经历了巨大的变动。根据路径的结果统计，新群建立会淘汰原有社群 80%左右的成员，而吸收新成员数约占新群总体人数的 50%左右。对淘汰成员的类型进行统计发现，90%以上的淘汰成员是原有群中的沉默围观者和疏离者，而加入到新群中的原有成员则主要为原有群的核心成员和活跃分子。因此，我们认为微信群间的演化包含一种特定的收缩机制，在新群生成时会主动过滤掉原有群内的非活跃成员，保留原有的活跃成员，同时吸收新鲜血液，进而完成资源的优化升级和重组分配。

为进一步探索微信群间的演化，我们对路径 QA-QC-QF-QG 上的用户发言内容进行文本分析，通过文本挖掘方法，抽取 4 个群内发言中的名词、动词等具有特定意义的词语，并过滤掉在前几个群中出现过的词，例如对 QA、QC 和 QF 的聊天记录进行分词、词性标注、去停用词、统计处理等，获取两个群记录的词表，如果词 w 在 QC 中出现，而没有包含在 QA 和 QC 的词表中，则将 w 保存下来， w 我们称之为 QF 产生新词。这样分析的原因在于，首先，用户间的聊天记录包含了大量的日常交流，而个体间的交流经常是最重要的信息资源^[36]，记录中有价值的观点和意见往往难以判别，而日常交流的常用词基本相似，因此，过滤后新词的出现往往预示着新观点的产生。其次，从信息资源的角度分析，新出现的词语往往能够代表新话题的产生，同样也代表了群内资源的生长。经过文本分析，结果如表 4 所示。

表 4 微信群新词统计表

微信群编号	QC	QF	QG
-------	----	----	----

包含词数	3486	1946	3926
新词数量	773	521	1233[L1]
更新比率	22.17%	26.77%	31.4%

从结果看来,新群在旧群的基础上增加了大约 20%到 30%左右的新词,[L2]根据我们之前的研究,核心成员的地位是在群的初始阶段建立的起来的,而后加入的成员往往会沦为沉默围观者和疏离者,随着时间的持续延长,群成员的活跃度下降,其提供资源的可能性和动力也随之减小[L3],因此,新群的开辟具有重要的资源更新和成员意识唤醒的价值。

6 总结

本研究基于真实的微信群数据,从群内和群间两个方面对微信群的演化规律进行了分析,主要研究发现包括以下几个方面:

首先,微信群的成立目的对于群的持续活跃时间存在直接的影响,基于活动与讨论的社群更易于引发群内互动,加深成员交流和建立强关联人际网络。从群的生长过程来看,微信群的演化过程可以划分为 6 个不同的阶段,包括发起、扩张、活跃、稳定、沉寂和分裂。

其次,在社群稳定阶段,微信群成员根据活跃程度可以划分为核心成员、活跃成员、半活跃成员、沉默围观者以及疏离者,其中大部分成员都是务实的利己主义者,成员间关系的建立是一种资源互补的关系。

再次,核心成员可能在群内形成一个较小的子群,但在总体网络中影响力较低,很难把持独立的话语权。核心成员的主要功能在于激发群成员的群体身份,构建社群的集体意识,其地位是在群建立的初始阶段形成的,而随着群的扩张和逐渐消寂,核心成员不再活跃,而是选择在其他新群中活动或主动创建新群。

最后,微信群间的演化包含一种特定的收缩机制,在新群生成时会主动过滤掉原有群内的非活跃成员,保留原有的活跃成员,吸收新鲜血液,进而完成资源的优化升级和重组分配,同时,新群的开辟具有重要的资源更新和成员意识唤醒的价值。已有的社群关系管理研究显示,通过成员推荐建立起一种持续的用户吸引力对于社群的持续发展具有重要作用^[37],但对于微信群来说,用户吸引力则并没有那么重要,随着群内资源的枯竭和不信任程度加重,旧群的衰退是一种必然的演化过程。

本研究同样存在几点局限和不足。第一,研究的案例数据背景较为单一,不能全面揭示各种类型群的演化规律,而且群内成员的人口学信息,包括职业、兴趣、教育背景、信息行为习惯等都影响着群的组成和生长。第二,忽视了情感在群[L4]内和群间演化过程中的作用,特别是对信任和冲突的关注不足。因此,进一步的研究需要更大的样本数据的支撑。第三,对成员关系的选择过于简单,只利用@关系构建网络减少了成员间隐含的内部交流数据,这就使得社群网络的生长过程过于单调,缺乏动态性的延伸。因此,更加深入的分析文本内容,从话题演化和情感关系的角度分析微信群的演化将是下一步的研究重点。

参考文献

- [1] Rheingold H. The virtual community: Homesteading on the electronic frontier[M]. MIT press, 1993.
- [2] Koh J, Kim Y G, Butler B, et al. Encouraging participation in virtual communities[J]. Communications of the ACM. 2007, 50(2): 69-74.
- [3] 党昊祺. 从传播学角度解构微信的信息传播模式[J]. 东南传播. 2012(7): 71-72.
- [4] 周蕾. 微信广告传播力研究[J]. 东南传播. 2012(1): 21-23.
- [5] 陈雅静. 微信,还能红多久?——以经济学SWOT理论分析微信[J]. 新闻知识. 2012(6): 83-84.
- [6] 方兴东,石现升,张笑容等. 微信传播机制与治理问题研究[J]. 现代传播. 2013, 35(6):

122-127.

- [7] 黄浩波,何卫华,叶青. 微信及其在图书馆信息服务中的应用[J]. 图书馆学报. 2013, 35(1): 62-64.
- [8] 肖金华,黄丽红. 基于微信的图书馆信息服务模式研究[J]. 现代情报. 2013, 33(6): 55-57.
- [9] 江波,覃燕梅. 基于微信的移动图书馆APP服务系统设计与实现[J]. 现代情报. 2013, 33(6): 41-44.
- [10] 杜文龙,徐光辉,冯现永. 高校微信用户利用行为实证分析与研究——以西安航空学院为例[J]. 图书馆学研究. 2015(3):67-70.
- [11] 张青青,刘佳音. 微信用户日常生活信息获取行为研究[J]. 情报探索. 2015(1): 75-78.
- [12] 马良. 国内微信用户软件使用行为影响因素研究[D]. 华中科技大学, 2012.
- [13] 陈佩. 微信主流用户的“使用与满足”研究[D]. 华中师范大学, 2014.
- [14] 刘鹏飞,何川. 舆情素养: 微信群发展问题与对策[N/OL]. 人民网, 2015-04-16[2015-06-11]. <http://jx.people.com.cn/n/2015/0416/c355236-24520832.html>.
- [15] 刘永丹. 基于Net logo的微信群舆情演化模型实现[J]. 无线互联科技. 2013(11): 197-198.
- [16] Tajfel H. Experimental Studies of Intergroup Behaviour[J]. Nato Advanced Study Institutes, 1982:227-246.
- [17] Bagozzi R P, Dholakia U M. Open Source Software User Communities: A Study of Participation in Linux User Groups[J]. Management Science. 2006, 52(7): 1099-1115.
- [18] Herrmann A, Dholakia U M, Algesheimer R. The social Influence of Brand Community: Evidence from European Car Clubs[J]. Journal of Marketing: A Quarterly Publication of the American Marketing Association. 2005, 69: 19-34.
- [19] Hogg M A, Hains S C. Intergroup relations and group solidarity: Effects of group identification and social beliefs on depersonalized attraction[J]. Journal of Personality and Social Psychology. 1996, 70(2): 295-309.
- [20] Asforth B E. Social identity theory and the organization[J]. Academy of Management Review. 1989, 14(1): 20-39.
- [21] Sen S, Bhattacharya C B. Consumer-Company Identification: A Framework for Understanding Consumers' Relationships with Companies[J]. Journal of Marketing: A Quarterly Publication of the American Marketing Association. 2003, 67(2): 76-88.
- [22] Aharne M. Antecedents and Consequences of Customer-Company Identification: Expanding the Role of Relationship Marketing[J]. Journal of Applied Psychology. 2005, 90(3): 574-585.
- [23] Mislove A, Lehmann S, Ahn Y Y, et al. Understanding the Demographics of Twitter Users[J]. Fifth International Aaai Conference on Weblogs & Social Media. 2011.
- [24] 刘瑛,杨伯淑. 互联网与虚拟社区[J]. 社会学研究. 2004(5): 1-7.
- [25] Dutton J E, Dukerich J M, Harquail C V. Organizational images and member identification[J]. Administrative Science Quarterly. 1994: 239-263.
- [26] Okleshen C, Grossbart S. Usenet groups, virtual community and consumer behaviors[J]. Advances in Consumer Research. 1998, 25(1): 276-282.
- [27] Bhattacharya C B, Rao H, Glynn M A. Understanding the bond of identification: An investigation of its correlates among art museum members[J]. The Journal of Marketing. 1995: 46-57.
- [28] 毛波,尤雯雯. 虚拟社区成员分类模型[J]. 清华大学学报: 自然科学版. 2006, 46: 1069-1073.
- [29] Wasko M M, Faraj S. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice[J]. MIS Quarterly. 2005: 35-57.
- [30] Wagner C, Prasarnphanich P. Innovating Collaborative Content Creation: The Role of Altruism

and Wiki Technology[J]. System Sciences, HICSS 40th Annual Hawaii International Conference on, 2007:18.

[31] Chen C J, Hung S W. To give or to receive? Factors influencing members' knowledge sharing and community promotion in professional virtual communities[J]. Information & Management. 2010, 47: 226-236.

[32] 艺术与传媒学院, 北师大传播效果实验室. 北师大发布首份微信社会资本对集体行为的影响报告[N/OL]. 北京师范大学官方网站, 2014-10-20[2015-06-11]. <http://www.bnu.edu.cn/xzdt/76597.html>.

[33] Muniz Jr A M, O Guinn T C. Brand community[J]. Journal of Consumer Research. 2001, 27(4): 412-432.

[34] Bagozzi R P, Dholakia U M. Intentional social action in virtual communities[J]. Journal of Interactive Marketing. 2002, 16(2): 2-21.

[35] Newman M E. Modularity and community structure in networks[J]. Proceedings of the National Academy of Sciences. 2006, 103(23): 8577-8582.

[36] Katona G, Mmuller E. A study of purchase decisions[C]. Consumer Behaviour: The Dynamics of Consumer Reaction, L. H. Clark (Ed.), New York: New York University Press, 1955, 30-87.

[37] Algesheimer R, Dholakia U M, Herrmann A. The social Influence of Brand Community: Evidence from European Car Clubs[J]. Journal of Marketing: A Quarterly Publication of the American Marketing Association. 2005, 69: 19-34.

[i] 丁元竹. 社区的基本理论与方法[M]. 北京: 北京师范大学出版社, 2009:69

[ii] Park, R. E., Burgess, E. W. and MacKenzie, R. D. The City[M]. Chicago: University of Chicago Press, 1925.