



在线健康社区中 用户回帖行为影响机理研究

刘璇¹, 汪林威¹, 李嘉¹, 张朋柱²

1 华东理工大学商学院, 上海 200237

2 上海交通大学安泰经济与管理学院, 上海 200052

摘要: 在线社交网络的持续发展和公民健康意识的不断增强促进了在线健康社区的兴起和繁荣。健康社区为用户提供了一个获取医疗资源并分享知识、经验和情感的开放式平台。由于其在为用户和患者提供丰富的医疗信息和知识资源、满足用户间的社交需求等方面表现突出, 近年来, 国内外在线健康社区都出现了蓬勃发展的态势。

在社区的各种行为中, 发帖和回帖是社区中用户社交行为的重要体现, 是维系社区繁荣的根本, 因此对社区中用户回帖行为的影响机理进行研究具有重要的意义。

以糖尿病在线健康社区甜蜜家园网站为研究对象, 基于该网站2015年1月~6月的社区发帖和回帖数据以及所有用户的个人信息数据, 构建健康社区的用户回帖有向网络, 该网络包含边和节点两部分属性信息; 运用指数随机图模型, 探讨该网络中网络结构和节点属性对回帖网络形成的影响机理。

研究表明, 回帖网络同时具有互惠性和传递性特征, 互惠性表明社区中的用户倾向于互相回帖, 传递性表明该网络有很好的发展潜力。节点属性对用户回帖行为的影响机制较为复杂, 其中, 用户倾向于对同质性(类型同质性)用户回帖; 社会资本(如好友数量、活跃度)高的用户能获得更高的回帖概率; 大量给他人回帖的用户和新用户更易获得他人的回帖。

研究结果进一步丰富了在线健康社区中用户行为机理的研究范式, 以及电子健康、在线社交网络、健康社区、论坛回帖行为等领域的研究, 并对社区的管理有一定的启示作用; 有利于指导社区的设计机制, 如鼓励新人多发帖、引导有较大影响力的用户(如管理者和活跃度高的用户)发帖互动、进一步促进同质用户间的信息交流等, 这些机制进一步促进和引导论坛中用户间的信息交互, 对论坛持续繁荣发挥重要作用。

关键词: 在线健康社区; 回帖行为机理; 指数随机图模型; 网络结构; 节点属性

中图分类号: G20 **文献标识码:** A **doi:** 10.3969/j.issn.1672-0334.2017.01.006

文章编号: 1672-0334(2017)01-0062-11

收稿日期: 2016-09-10 **修返日期:** 2016-12-28

基金项目: 国家自然科学基金(71371005, 71471064, 91646205); 上海市浦江人才计划项目(15PJC019); 中央高校基本科研业务费(WN1522008)

作者简介: 刘璇, 管理学博士, 华东理工大学商学院副教授, 研究方向为知识管理、电子商务和大数据决策等, 代表性学术成果为“Determining inventor status and its effect on knowledge diffusion: a study on nanotechnology literature from China, Russia, and India”, 发表在2011年第6期《Journal of the American Society for Information Science and Technology》, E-mail: xuanliu@ecust.edu.cn

汪林威, 华东理工大学商学院硕士研究生, 研究方向为电子健康和社交网络分析等, E-mail: leeven92@163.com

李嘉, 管理学博士, 华东理工大学商学院副教授, 研究方向为决策支持系统、电子商务和知识管理等, 代表性学术成果为“支持模糊型任务的信息组织结构研究”, 发表在2015年第5期《管理科学学报》, E-mail: jiali@ecust.edu.cn

张朋柱, 管理学博士, 上海交通大学安泰经济与管理学院教授, 研究方向为电子商务、电子政府和金融信息系统等, 代表性学术成果为“G2G information sharing among government agencies”, 发表在2014年第1期《Information & Management》, E-mail: pzzhang@sjtu.edu.cn

引言

电子健康(e-health)是信息和通信技术在医疗卫生领域从预防、诊断、治疗、随诊、康复到健康促进全方位的应用,是信息技术、医学、公共卫生、管理学和社会学的交叉学科,其目的是最大限度地整合和利用医疗健康资源,提高公众的健康状况^[1-2]。电子健康改变着人们的健康生活,也开创和促进了很多新的技术和服务理念,从而对传统医疗行业带来深刻的变革,这些变革包括但不仅限于覆盖全民的健康信息网络、电子健康记录、远程医疗服务、移动医疗设备和通信等,越来越多基于IT和通信技术的疾病预防、诊断、治疗、健康监测和生活方式管理的系统和设备,以及基于社会化媒体的医疗活动,互联网逐渐成为公众获取健康信息、寻求医疗保健活动的重要资源^[3-4]。

与传统的电子病历相比,社会化媒体是电子健康一个新兴的应用领域。在线健康社区(online health communities, OHCs)是在线社交网络持续发展与公民健康意识不断增强相融合的产物,是指基于web 2.0互联网技术支持,用户针对疾病问题进行治疗方案或病情等相关知识经验的共享、专家咨询和成员交流等活动的在线社区^[5]。在线健康社区中,人们可以向专家咨询问题,向其他病患者学习或与其他病人一起分享各自的经历,获取有益的健康知识^[6]。近年来,Patientslikeme、CureTogether、DailyStrength、Disaboom、甜蜜家园、天涯社区健康在线、39健康网等国内外在线健康社区都出现了蓬勃发展的态势。在线健康社区中,发帖和回帖是用户社交行为的重要表现方式,是维系社区繁荣的根本,因此对社区中用户回帖行为的影响机理进行研究具有重要的意义。

1 相关研究评述

回帖行为是在线社区中一种用户间信息交互的过程,学者们对网络社区的主要特征和传播模式做了大量的研究。随机指数图模型^[7]、小世界模型^[8]和无标度模型^[9]是3种解释网络拓扑特征产生的原因和演化模式的基本网络,在这些网络模型的基础上,已有研究对不同的网络社区进行实证研究。ADAMIC et al.^[10]研究Stanford大学的一个互联网社区,发现人际网络拓扑结构具有小世界特征;AHN et al.^[11]对在线互联网社区Cyworld进行度分布、聚类系数、度耦合、平均路径长度等网络特征分析,证明人际关系网络具有小世界和幂律分布特性;GIRVAN et al.^[12]的研究发现网络用户在互联网社区中倾向于形成具有强链接的社团结构。

用户的回帖行为受到众多因素影响,已有研究从社区参与动机和影响因素角度讨论在线社区用户参与行为。陈然^[13]从参与动因角度研究社会民生网络社区中参与者发帖和回帖的动因,主要包括参与者的个体因素和社会因素两方面,个体因素又包括外在动机(即感知有用性)和内在动机(即感知愉悦),外在动机分为工具性动机、逃避现实性动机、

匿名表达动机和声望动机4类,工具性动机包括社会支持动机和信息交流动机,社会因素主要指参与者的社区认同感;研究发现社区认同感、信息交流动机和声望动机是促使参与者围绕社会民生话题发表主帖的显著动因,社区认同感和社会支持动机是促使参与者围绕社会民生话题发表回帖的重要推动力。已有关于在线社区参与的影响因素的研究主要集中在角色和内容两大方面。BENOIT et al.^[14]通过测试用户参与在线社区的相关行为,发现用户对自我角色的清晰认知和执行、社区提供者的积极响应、用户对社区的满意度都会影响用户的一系列参与行为;丁汉青等^[15]认为社交网络中成员的影响力与知识位势之间存在一定的关联性,即社交网络的用户之间不仅有以知识形式存在的信息流,也存在用户之间的影响流;而信息的发布、回复、评论过程中,影响力总是倾向于从信息位较高的用户流向信息位较低的用户,影响力在社交网络中往往通过较高的发言质量和频率等形式衡量。

已有对在线健康社区的研究主要体现在用户参与动机、群组特征和主题特征几个方面。在用户参与动机方面,KUMMERVOLD et al.^[16]研究发现,大部分女性认为参与健康社区有助于她们获取医疗信息、建立更多的社会关系和得到更强的社会支持,因此多数人更愿意在线上谈论私人健康问题,而不是面对面交谈,甚至差不多一半的人只在线上谈论她们的健康问题。在群组特征方面,有学者研究不同群组的特征对用户参与在线社区行为的影响。VYDISWARAN et al.^[17]对健康社区中建立小群组进行研究,发现最初具有同质性的用户会促成他们形成小组,有更多的社交活动;WU et al.^[18]分析了影响在线健康社区中用户沟通交流的因素,发现在线健康社区中用户的发帖与回帖有明显的互惠关系,拥有较高的活跃度的用户群对社区体验有较高的满意度,同时拥有更丰富的经历和偏激的情感的用户组发的帖子有更重要的影响。还有学者对在线社区中主题信息相关特征进行研究,SUDAU et al.^[19]发现在线健康社区中大量对人们有影响的信息不是来自科学研究,而是社交媒体资源;张星等^[20]的研究表明,论据质量、信息完整性、表达质量、一致性和来源可信性对在线健康社区的信息可靠性均有显著影响;金碧漪等^[21]以糖尿病为例,以来自健康社区的社会化标签社区和社会化问答社区的问答记录作为研究对象,选取八大类主题,比较两种网络社区中该八大主题分布情况的异同,促使各网络平台能够更好地提供在线健康信息服务。

研究者们虽然对互联网社区中的社交网络关系和用户参与行为开展了大量的研究,对其幂律分布、小世界行为等结构特性达成了部分共识,但是对具体的网络结构及其形成原因和影响并未取得一致结论。另外,已有研究发现用户特征与社区网络用户参与行为之间的关联,但少有研究全面考虑不同用户特征的作用效果和不同特征的不同影响机制,对

用户属性和网络特征如何交互影响社区用户的参与更是少有涉及。基于此,本研究对这两个方面进行更加具体的探讨。

社区中有发帖者和回复者,发帖者的第一个发文即为帖子的主题,在主题后面的留言则为回文,社区中的用户回帖网络是指用户给主题帖留言形成的网络系统,该网络中节点为发帖者和回帖者,一条有向边是指一位回帖者给发帖者一次留言,以用户为节点形成的网络系统是目前研究的主要形式^[22]。本研究从回帖网络的网络结构特征和社区中用户的节点属性两个角度探讨二者对回帖网络形成的影响机理,采用指数随机图模型检验这两个方面对用户回帖关系形成的影响。指数随机图模型是一种用来分析原网络中的结构是否呈现出我们假设的趋势的统计模型,它能同时包含并处理复杂的网络结构和多重节点属性^[23-24]。

2 理论基础和研究假设

为了验证社区回帖网络的网络结构和节点属性对回帖网络形成的影响,本研究基于前人的理论提出一系列假设。

互相回帖是互惠关系的一种体现^[25],节点*i*回复了节点*j*的帖子,随后节点*j*在一段时间内至少回复节点*i*一次,这就形成了互相回复关系,这也是用户之间充分交互的体现^[26],健康社区中用户间的互惠性回帖是社会交换的一种体现,有利于促进用户间的知识交换和情感交互。因此,本研究提出假设。

H₁ 健康社区中用户间倾向于互相回帖。

网络传递性是社交网络一个重要的性质^[27]。在社交网络中包含多种网络结构,可分为有向网络和无向网络,实体由节点来表示,关系由边来表示。传递性为社交网络中存在三元关系的程度。根据KLYMKO et al.^[28]的研究,有向网络中三元关系有7种形式,见图1,它们都体现了网络的传递性。社交网络传递性不仅反映存在于3个行动主体之间的三元路径,还反映在整个网络中将要发生的结构型事件^[29]。在线健康社区中回帖网络是一种有向网络,

用户间达到信息传递和信息扩散的目的。因此,本研究提出假设。

H₂ 回帖网络具有传递性。

同质性是社交网络一个重要的维度。MCPHERSON et al.^[30]提出社交网络中的同质性原则,即两个相似的人建立联系的概率比那些相似性较低的人之间建立联系的概率要高得多。个体间的同质性可以用人口统计学的各个方面测量,如种族同质性、年龄同质性、宗教同质性、性别同质性、婚姻状况同质性等^[31]。SONG et al.^[32]将同质性理论运用到电子健康的研究中,研究治疗方式和健康状况的同质性对病人之间好友关系形成的影响。在健康社区中,基于同质性理论,探讨回帖网络中用户属性(包括性别、社区类型、用户地域)的同质性是否促进回帖关系的形成。因此,本研究提出假设。

H_{3a} 用户倾向于对相同性别的用户回帖;

H_{3b} 社区中相同类型的用户之间更容易形成回帖关系;

H_{3c} 相同地域的用户之间有更多的回帖关系。

除个体间人口统计学属性外,社交网络的形成还受到处于社交网络中每个个体的社交属性的影响,用户的社交性使用户对关系的形成拥有偏好性^[33]。社会资本理论表明,拥有较高的关系型社会资本(以下用好友数和活跃度衡量)的用户,将有更高可能在社交网络中获得更多的社会资源。因此,本研究提出假设。

H_{4a} 拥有较多好友的用户会得到更多其他用户的回帖;

H_{4b} 活跃度高的用户会得到更多其他用户的回帖。

社交网络也是知识传递的途径^[34],ARGOTE et al.^[35]定义知识转移是一个网络成员受到其他成员经验影响的过程,强调网络里知识传递的发生主要是源于他人的影响。健康论坛的发帖-回帖关系,体现的是一种社会交换的过程,知识和情感这两种典型的社会资源得以通过发帖-回帖关系进行交换,在社会交换的过程中,拥有较高的结构型社会资本

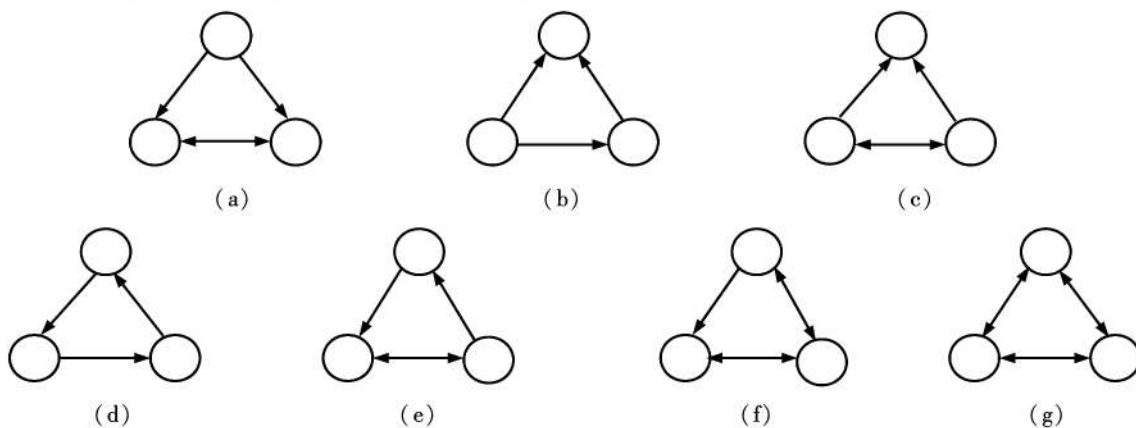


图1 有向网络三元结构

Figure 1 Triangles in Directed Networks

表1 用户属性
Table 1 User Attributions

用户属性	对应假设	变量类型	含义	测量方法
性别	H _{3a}	二元分类变量	用户的性别	女性取值为0,男性取值为1
社区类型	H _{3b}	分类变量	用户在社区中的身份识别	用户类型包括医生、家属、客服、I型糖尿病、II型糖尿病,每种类型为一类
用户地域	H _{3c}	分类变量	每个用户的所在地	每个省为一类
用户好友数	H _{4a}	二元分类变量	用户的社交能力	有较多好友的用户,好友数量多的前25%的用户取值为1,其他取值为0
用户积分	H _{4b}	二元分类变量	用户的活跃度	积分最高的前25%的用户取值为1,其他取值为0
用户回帖数量	H _{5a}	二元分类变量	回帖网络中用户节点的出度	出度数高的前25%的用户取值为1,其他取值为0
用户注册时间	H _{5b}	二元分类变量	表示新或老用户	注册时间最新的25%的用户取值为1,其他取值为0

(用回帖数量测量)的用户,由于对社区贡献度高,更容易得到回馈。因此,本研究提出假设。

H_{5a} 大量给别人回帖的用户也会得到更多其他人的回帖。

另外,拥有高势位的用户往往拥有高质量的知识产物,知识的传递也从他们流向低势位的用户^[36]。在线健康社区中,新用户是一类知识势位较低的用户,因此,本研究提出假设。

H_{5b} 新用户会得到更多的回帖。

3 数据和方法

3.1 数据描述

本研究从糖尿病社区甜蜜家园(<http://bbs.tmbz.com>)收集二手数据。甜蜜家园是目前中国知名的糖尿病在线社区,社区中的用户包括糖尿病病人、医生、家属和客服等。该社区创办于2005年,已拥有大约180 000名用户,社区给病人提供了一个分享治疗心得、寻求帮助的平台,给家属提供了一个情感支持和出谋划策的场所,给医生和客服提供了一个帮助他人、提供先进医疗资源的平台。在当前的医疗水平下,糖尿病依旧是一种无法完全根治的慢性疾病,这就要求糖尿病病人持续注意自己的健康,不断关注最新的医疗成果,通过社区获得知识共享和情感交互。由于社区中用户发帖-回帖行为是用户社交活动的重要组成部分,有助于提高用户粘性,保持社区用户活跃度,促进平台的发展和繁荣,这也是本研究选取糖尿病社区研究用户发帖-回帖行为的原因。

本研究收集2015年1月~6月的所有发帖和回帖数据以及相关用户的所有用户属性数据,初始数据包括2 465名用户、15 297条发帖和69 394条回帖。根据假设,本研究需要构建回帖网络,并实证分析网络结构和个体属性对回帖关系的影响。因此,对数据

进行筛选,最终得到属性完整的用户277名,构建了这277名用户之间的回帖网络,共872条边。

表1给出用户属性的描述性统计。用户属性中,性别为二元分类变量,0为女性,1为男性;社区类型为分类变量,每个类型为一类;用户地域为分类变量,每个省为一类;用户好友数为连续变量,是每个用户好友的数量;用户积分为连续变量,根据论坛的说明,用户总积分=精华帖数×10+威望×50+金钱,此处威望包含用户的精华贴数和排行榜排名,金钱通过发帖和回帖的方式获得,总体来说,积分能表征用户的发帖和回帖积极性,也表征用户的权威性,作为一个综合性指标,用以表示用户的活跃度;用户回帖数量为连续变量,表征回帖网络中用户节点的出度大小;用户注册时间为连续变量,用以表示新老用户。为更好地契合和检验假设,根据已有研究^[37-38],本研究将用户好友数、用户积分、用户回帖数量、用户注册时间4个连续变量处理为0-1分类变量。表2给出经过处理后用于测试的用户属性数据的均值和标准差,其中,性别均值为0.668,表明男性

表2 属性均值和标准差
Table 2 Mean and Standard Deviation of Attributions

用户属性	均值	标准差
性别	0.668	0.471
用户好友数	0.249	0.432
用户积分	0.249	0.432
用户回帖数量	0.249	0.432
用户注册时间	0.249	0.432

表3 相关系数
Table 3 Correlation Coefficient

		用户好友数	用户积分	用户回帖数量	用户注册时间
用户好友数	Pearson 相关性	1	0.709**	0.183**	0.280**
	显著性(双侧)		0	0.002	0
用户积分	Pearson 相关性	0.709**	1	0.180**	0.396**
	显著性(双侧)	0		0.003	0
用户回帖数量	Pearson 相关性	0.183**	0.180**	1	-0.026
	显著性(双侧)	0.002	0.003		0.663
用户注册时间	Pearson 相关性	0.280**	0.396**	-0.026	1
	显著性(双侧)	0	0	0.663	

注:样本数为277,**为 $p < 0.010$ 。下同。

多于女性,标准差为0.471。其他属性均处理为二元分类变量,选取前25%取值为1,其余取值为0,因此拥有相同的均值0.249和标准差0.432。

进一步地,属性之间的相关系数见表3。由表3可知,用户的好友数和积分两个属性之间有很强的相关性,因此本研究在后文的稳健性检验中将分别去掉其中一个属性,检验模型的稳健性。其他属性之间的相关性并不强,可以同时进入检验模型。

为了更好地识别实验数据的代表性,本研究对实验抽取的277名用户的属性与原始数据中2465名用户的属性进行对比,表4给出分类变量各类别所占比例和连续变量的均值。由于社区中大量用户对性别、地域等相关信息选择保密,出现实验数据占比偏低的情况。但对比实验数据与原始数据发现,男性和女性的比例、用户地域集中的省份基本保持一致。此外,4个连续变量中,用户积分和用户回帖数量的均值其原始数据与实验数据存在一定偏差,后续本研究对4个变量调整分类标准,进行模型的稳健性检验。

本研究依据实验数据构建回帖网络,网络结构见图2。在由277个节点组成的有向网络图中,含有872条有向边,网络密度为0.011,二元交互网络有40对,三元结构有1464个。

3.2 指数随机图模型

本研究运用指数随机图模型(exponential random graph model, ERGM)检验假设。ERGM是一种能够同时融合并分析网络结构和节点属性的统计方法,该模型将网络关系图转变为数学模型,在网络图中, n 为节点数,随机变量 $Y_{i,j}$ 为 i 节点与 j 节点之间的联系, $Y_{i,j} = 1$ 时表示 i 节点与 j 节点之间存在联系,否则表示不存在联系^[39]。ERGM假设网络由随机过程产生,ERGM根据原有的网络产生随机的网络图,然后比较产生的网络与原有的网络,它们越相似,则模型模拟

表4 实验用户与原始用户属性对比
Table 4 User Attributions Comparison between Experimental Users and Original Users

用户属性	对比角度	原始数据	实验数据
性别	男性/%	24.138	66.787
	女性/%	13.834	33.213
	保密/%	62.028	0
社区类型	I型糖尿病/%	19.757	22.744
	II型糖尿病/%	68.235	70.397
	其他/%	11.968	6.859
	未填/%		
用户地域	广东/%	7.343	12.274
	江苏/%	5.842	9.025
	北京/%	5.152	5.776
	山东/%	4.381	6.137
	浙江/%	3.529	9.025
	辽宁/%	3.043	6.859
	其他/%	34.077	50.904
	未填/%	36.633	0
用户积分	均值	737	1935.947
用户好友数	均值	2.343	4.481
用户回帖数量	均值	13.978	3.150
用户注册时间	均值	2.674	3.400

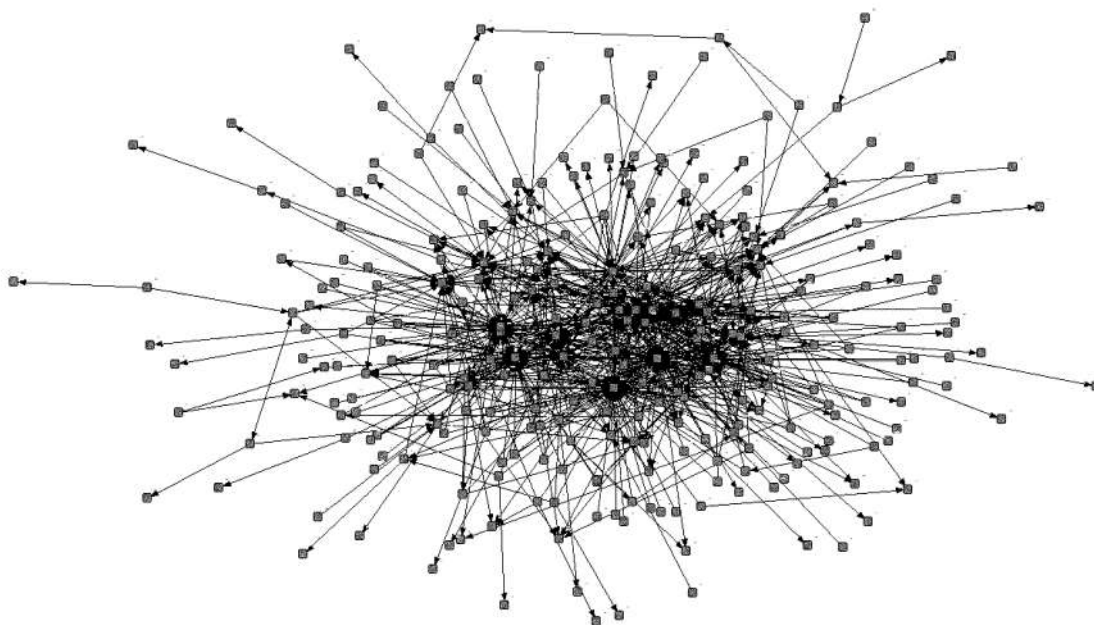


图2 回帖网络

Figure 2 Reply Networks

得越好。最终模型估计出各种结构的参数,该参数表示其对应的网络结构在网络图中出现的频率^[40]。ERGM的一般数学表达式为

$$\Pr(X = x) = \frac{1}{k} \exp \left[\sum_A \theta_A g_A(x) \right] \quad (1)$$

其中, X 为利用模型生成的网络; x 为真实的网络,称为观测网络; k 为一个归一化参量,统计并算出所有概率取值,用以确保(1)式为一个正确的概率分布^[41]; A 为观测网络中所有的网络结构的集合; θ_A 为对应于网络结构 A 的参数; $g_A(x)$ 为对应于网络结构 A 的网络统计值,当网络结构 A 在网络 x 中出现时 $g_A(x) = 1$,否则 $g_A(x) = 0$ 。本研究的目的是利用观测网络(包括图的有向连接和节点属性等)分析并求出模型的最佳参数,使生成的模型网络与观测网络最接近。

为详细说明ERGM如何运用到检验本研究假设中,本研究列出每个假设对应的分析层级和网络结构,表5给出所有假设以及对应的分析层级和网络结构的图解。

4 结果和讨论

表6给出ERGM参数估计的结果,根据已有研究,判断估计得到的参数结果是否显著,需要检验该参数是否至少为标准差的两倍^[42]。

由表6可知,检验 H_1 和 H_2 的参数分别为2.187和3.626,两个正的较大的系数表明这两种网络结构出现的概率很大,即社区中的用户倾向于互相回帖,且回帖网络具有很强的传递性, H_1 和 H_2 得到验证。表明健康社区中存在着用户之间的互惠性,这有助于促进知识交互和情感交流,网络的传递性促进了用户之间更多的回帖和更充分的信息扩散,也是保证

社区持续繁荣的根本所在。

在节点属性方面,由表6可知,检验 H_{3a} 的参数为-5.909,表明同性之间互相回帖的趋势不明显,有可能是因为异性之间更具有吸引力,导致相同性别这一因素对回帖无影响, H_{3a} 没有得到验证。这与已有研究一致^[32]。检验 H_{3b} 的参数为0.724,且结果显著, H_{3b} 得到验证;并由 $\exp(0.724) \approx 2.063$ 可知,相同社区类型的用户之间多出一倍的回帖概率,即用户的社区类型同质性能促进用户之间的回帖。表明用户倾向于与自己有相似身份或疾病类型的人交流互动,因为用户能从有相似身份或疾病类型的人身上得到更多的社会认同和信息交流。检验 H_{3c} 的结果不显著, H_{3c} 没有得到验证,说明用户地域上的一致对用户回帖选择没有显著影响,在线社交网络的一大特征就是用户可以突破地域界限进行社交活动,因而地域上的同质性对回帖网络的形成影响并不显著。检验 H_{4a} 的参数为0.807,且结果显著, H_{4a} 得到验证;由 $\exp(0.807) \approx 2.241$ 可知,拥有较多好友的用户会得到两倍多于其他用户的回帖。好友数量多是用户交际能力的体现,他们在社交网络中具有核心地位,其发帖容易获得更多的关注,获得大量好友甚至陌生人的回应。检验 H_{4b} 的参数为1.699,且结果显著, H_{4b} 得到验证;由 $\exp(1.699) \approx 5.468$ 可知,活跃度高的用户显著可以得到大量的回帖,该节点属性对回帖网络的形成有很大的影响。活跃度高的用户在社交网络中一般处于核心地位,成为社区的焦点和核心,因而其发帖能吸引更多人关注。检验 H_{5a} 的参数为0.685,且结果显著, H_{5a} 得到验证,表明大量回复别人帖子的用户,其发帖也能得到近两倍($\exp(0.685) \approx 1.984$)多的回帖,这类用户通过大量地传递知识给其他用户,根据社会交换原理,其发帖请求也更容易

表5 因素类型、研究假设、分析层级和网络结构图解
Table 5 Factor Type, Research Hypotheses, Levels of Analysis, and Graphical Illustrations for Network Configurations

因素类型	假设	分析层级	图解
网络结构	H ₁ 健康社区中用户间倾向于互相回帖	二元层级	节点 A $\xleftrightarrow{\text{回帖关系}}$ 节点 B
	H ₂ 回帖网络具有传递性	三元层级	节点 A $\xleftrightarrow{\text{回帖关系}}$ 节点 B 节点 A $\xleftrightarrow{\text{回帖关系}}$ 节点 C
用户属性	H _{3a} 用户倾向于对相同性别的用户回帖	二元层级	节点 A $\xleftrightarrow{\text{回帖关系}}$ 性别同 A 的节点 B
	H _{3b} 社区中相同类型的用户之间更容易形成回帖关系	二元层级	节点 A $\xleftrightarrow{\text{回帖关系}}$ 社区类型同 A 的节点 B
	H _{3c} 相同地域的用户之间有更多的回帖关系	二元层级	节点 A $\xleftrightarrow{\text{回帖关系}}$ 用户地域同 A 的节点 B
	H _{4a} 拥有较多好友的用户会得到更多其他用户的回帖	节点层级	其他节点 $\xleftrightarrow{\text{回帖关系}}$ 用户好友数多的节点
	H _{4b} 活跃度高的用户会得到更多其他用户的回帖	节点层级	其他节点 $\xleftrightarrow{\text{回帖关系}}$ 用户积分高的节点
	H _{5a} 大量给别人回帖的用户也会得到更多其他人的回帖	节点层级	其他节点 $\xleftrightarrow{\text{回帖关系}}$ 用户回帖数量多的节点
	H _{5b} 新用户会得到更多的回帖	节点层级	其他节点 $\xleftrightarrow{\text{回帖关系}}$ 用户注册时间短的节点

表6 ERGM 估计结果
Table 6 Results for ERGM Estimation

类型	假设	参数	标准差	p 值	结果
网络结构	H ₁	2.187	0.190	< 1e-04	支持
	H ₂	3.626	0.105	< 1e-04	支持
用户属性	H _{3a}	-5.909	0.101	< 1e-04	不支持
	H _{3b}	0.724	0.075	< 1e-04	支持
	H _{3c}	0.059	0.144	0.872	不支持
	H _{4a}	0.807	0.096	< 1e-04	支持
	H _{4b}	1.699	0.111	< 1e-04	支持
	H _{5a}	0.685	0.063	< 1e-04	支持
H _{5b}	0.447	0.090	< 1e-04	支持	

得到其他人的回馈。检验H_{5b}的参数为0.477,且结果显著,H_{5b}得到验证;由exp(0.447)≈1.564可知,注册时间短的新人发帖,比其他人发帖增加了56%的回帖概率,即新用户会得到更多的回帖关注。原因可能在于新用户的帖子一般出于求助和咨询的目的,或是期望寻求好友和分享情感,而新用户一般处于知

识和情感的低位势,因而能更多的从其他高位势用户中获取信息传递或者情感支持。

本研究通过6个模型检验原模型(M₀)的稳健性,模型1(M₁)和模型2(M₂)分别为去掉用户好友数和用户积分两个属性后得到的模型,模型3(M₃)~模型6(M₆)分别对应选取每个属性的前15%和前40%的用户为用户好友数多的、用户积分高的、用户回帖数量多的和用户注册时间短时模拟得到的结果,其他变量维持不变。稳健性检验结果见表7。

在表7中,由M₁和M₂可知,分别去掉用户好友数和用户积分后,本研究得到的结果和原模型M₀基本一致。另外,由M₃~M₆可知,在H₄和H₅在调整分类标准后,结果仅存在小幅度浮动,由此说明虽然实验数据的用户积分、用户好友数、用户回帖数量和用户注册时间较原始数据存在一定的偏移,但是这些变量对用户回帖关系构建的影响并不大。值得注意的是M₆中,H_{5b}在选取前15%的用户为新用户时,得到的参数为负数,且结果不显著,考虑到15%意味着选出的新用户数量偏少,导致网络结构出现频率较低,产生这样的结果合理。因此,通过对比M₁~M₆与M₀,本研究认为原始模型的结果拥有较好的稳健性。

为了检验模型是否较好的拟合实际网络数据,本研究给出拟合优度图,见图3。图3中呈现的是原有

表7 模型稳健性检验结果
Table 7 Results of Model Robustness Test

假设	含义	M ₀	M ₁	M ₂	M ₃	M ₄	M ₅	M ₆
		原模型	去掉用户 好友数	去掉用户 积分	H _{4a} (15%,40%)	H _{4b} (15%,40%)	H _{5a} (15%,40%)	H _{5b} (15%,40%)
H ₁	互惠性	2.187**	2.192**	2.184**	(2.181**, 2.177**)	(2.189**, 2.171**)	(2.196**, 2.187**)	(2.192**, 2.170**)
H ₂	传递性	3.626**	3.584**	2.133**	(3.631**, 3.598**)	(2.892**, 2.793**)	(3.345**, 3.132**)	(3.553**, 3.182**)
H _{3a}	性别	-5.909**	-6.150**	-5.619**	(-5.956**, -6.440**)	(-6.595**, -6.307**)	(-6.080**, -6.102**)	(-6.075**, -6.154**)
H _{3b}	社区类型	0.724**	0.784**	0.767**	(0.653**, 0.732**)	(0.743**, 0.708**)	(0.712**, 0.733**)	(0.685**, 0.688**)
H _{3c}	用户地域	0.059	0.023	0.030	(0.020, 0.081)	(0.022, 0.007)	(0.012, 0.019)	(0.040, 0.035)
H _{4a}	用户好友数 (25%)	0.807**		1.731**	(0.451**, 1.083**)	(0.706**, 0.679**)	(0.745**, 0.733**)	(0.653**, 0.704**)
H _{4b}	用户积分 (25%)	1.699**	1.470**		(1.784**, 1.762*)	(1.331**, 1.182**)	(1.452**, 1.392**)	(1.298**, 1.343**)
H _{5a}	用户回帖 数量(25%)	0.685**	0.663**	0.450**	(1.581**, 1.471**)	(1.419**, 1.458**)	(0.866**, 0.595**)	(1.366**, 1.354**)
H _{5b}	用户注册 时间(25%)	0.447**	0.459**	0.142*	(0.276*, 0.113*)	(0.083, 0.333*)	(0.343*, 0.255*)	(-0.254, 0.632**)

注:*为p < 0.050。

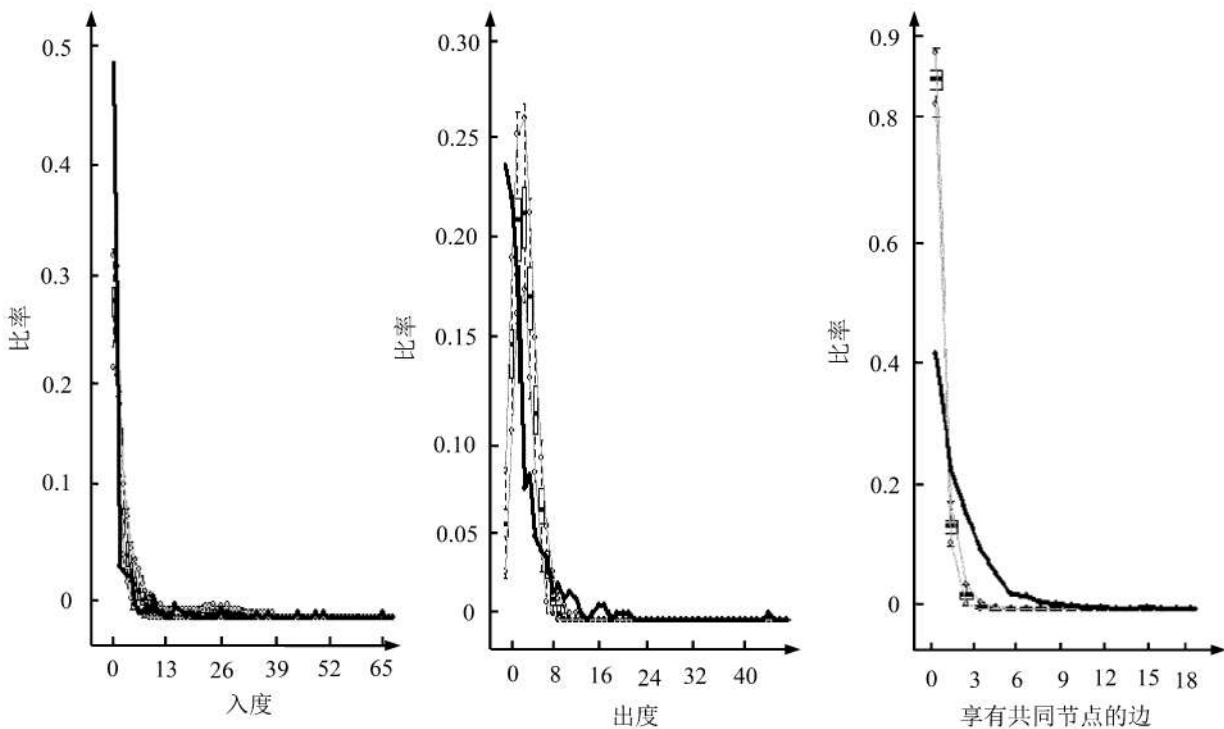


图3 拟合优度图
Figure 3 Goodness-of-fit Plot

网络参数和ERGM估计出的参数的比较,包括入度、出度和享有共同节点的边,实线代表原有的网络,虚线代表方差,箱形图代表模拟的网络参数。由图3可知,总体来说,估计参数与原网络差别较小,说明现有的ERGM模型拟合得较好。

5 结论

本研究选取在线健康社区甜蜜家园的回帖数据和用户属性数据,经过筛选数据和网络构建,形成用户间回帖有向网络,并利用ERGM分析网络结构和节点属性对回帖网络形成的影响,实证检验本研究提出的9个假设。研究结果表明,回帖网络同时具有互惠性和传递性特征;用户属性特征中,用户在社区中的类型同质性对回帖网络形成发挥积极的影响;好友数量多、活跃度高和新加入的用户能得到更高概率的回帖。

本研究融合并探讨网络结构和节点属性两个角度的影响,对于节点属性的研究进行多角度的分类,从同质性、知识传递性和社会资本多个角度处理属性变量,丰富了电子健康、在线社交网络、在线健康社区和论坛回帖行为等方面的研究。同时,研究结果能给在线健康社区提供管理上的参考建议,如鼓励新人以及管理者和活跃度高等有较大影响力的用户多发帖、多互动,有利于提高社区回帖数量和质量,服务和吸引新用户,提高老用户的活跃度和粘性等,最终能够促进社区繁荣,使更多患者从在线健康社区受惠。

本研究仅考虑单一的发帖-回帖网络,考虑到好友网络、用户访问网络与回帖网络可能存在很大的相关性,探索多重网络间的关系以及不同网络结构和属性的相互作用也具有重要的理论和现实意义;除网络结构和用户属性,帖子的内容对回帖网络的形成和影响也具有重要影响。这些工作将留待未来进一步深入研究。

参考文献:

- [1] EYSENBACH G. What is e-health?. *Journal of Medical Internet Research*, 2001, 3(2): e20-1-e20-2.
- [2] WICKS P, STAMFORD J, GROOTENHUIS M A, et al. Innovations in e-health. *Quality of Life Research*, 2014, 23(1): 195-203.
- [3] NORMAN C D, SKINNER H A. eHealth literacy: essential skills for consumer health in a networked world. *Journal of Medical Internet Research*, 2006, 8(2): e9.
- [4] SAPORITO B. The e-health revolution. *Time*, 2005, 165(26): 55-57.
- [5] DEMIRIS G. The diffusion of virtual communities in health care: concepts and challenges. *Patient Education and Counseling*, 2006, 62(2): 178-188.
- [6] 周军杰. 用户在线参与的行为类型: 基于在线健康社区的质性分析. *管理案例研究与评论*, 2016, 9(2): 173-184.
ZHOU Junjie. Examining the users' participation behavior types in online health communities: a qualitative approach. *Journal of Management Case Studies*, 2016, 9(2): 173-184. (in Chinese)
- [7] ERDÖS P, RÉNYI A. On the evolution of random graphs // *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, 1960: 17-61.
- [8] WATTS D J, STROGATZ S H. Collective dynamics of 'small-world' networks. *Nature*, 1998, 393: 440-442.
- [9] BARABÁSI A L, ALBERT R. Emergence of scaling in random networks. *Science*, 1999, 286(5439): 509-512.
- [10] ADAMIC L, BUYUKKOKTEN O, ADAR E. A social network caught in the Web. *First Monday*, 2003, 8(6): 1-22.
- [11] AHN Y Y, HAN S, KWAK H, et al. Analysis of topological characteristics of huge online social networking services // *Proceedings of the 16th International Conference on World Wide Web*. New York: ACM Publications, 2007: 835-844.
- [12] GIRVAN M, NEWMAN M E J. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 2002, 99(12): 7821-7826.
- [13] 陈然. 社会民生网络论坛参与者发帖行为的动因探讨: 基于五大中文网络社区的实证研究. *新疆社会科学*, 2012(6): 108-114.
CHEN Ran. Survey on the influencing factors to posting of participants in social livelihood online forum: an empirical study on five Chinese online communities. *Social Sciences in Xinjiang*, 2012(6): 108-114. (in Chinese)
- [14] BENOIT S, BILSTEIN N, HOGREVE J, et al. Explaining social exchanges in information-based online communities (IBOCs). *Journal of Service Management*, 2016, 27(4): 460-480.
- [15] 丁汉青, 王亚萍. SNS网络空间中“意见领袖”特征的分析: 以豆瓣网为例. *新闻与传播研究*, 2010, 17(3): 82-91.
DING Hanqing, WANG Yaping. Analyzing on 'opinion leader' attributes in SNS cyberspace: an investigation of douban.com. *Journalism & Communication*, 2010, 17(3): 82-91. (in Chinese)
- [16] KUMMERVOLD P E, GAMMON D, BERGVIK S, et al. Social support in a wired world: use of online mental health forums in Norway. *Nordic Journal of Psychiatry*, 2002, 56(1): 59-65.
- [17] VYDISWARAN V G V, LIU Y, ZHENG K, et al. User-created groups in health forums: what makes them special? // *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*. Ann Arbor, Michigan: AAAI Press, 2014: 515-524.
- [18] WU B, JIANG S, CHEN H. Effects of individuals' motivations on communications in online health forums. *Social Behavior and Personality*, 2016, 44(2): 299-312.
- [19] SUDAU F, FRIEDE T, GRABOWSKI J, et al. Sources of information and behavioral patterns in online health forums: observational study. *Journal of Medical Internet Research*, 2014, 16(1): e10-1-e10-16.
- [20] 张星, 夏火松, 陈星, 等. 在线健康社区中信息可信性的影响因素研究. *图书情报工作*, 2015, 59(22): 88-96, 104.

- ZHANG Xing, XIA Huosong, CHEN Xing, et al. A study of factors affecting the information credibility of online health communities from an elaboration likelihood model perspective. *Library and Information Service*, 2015, 59(22):88-96,104. (in Chinese)
- [21] 金碧漪, 许鑫. 网络健康社区中的主题特征研究. *图书情报工作*, 2015, 59(12):100-105.
JIN Biyi, XU Xin. Research on theme features in online health community. *Library and Information Service*, 2015, 59(12):100-105. (in Chinese)
- [22] 叶作亮, 王雪乔, 王仙玲, 等. 一类BBS网络统计特性实证分析. *复杂系统与复杂性科学*, 2010, 7(1):52-58.
YE Zuoliang, WANG Xueqiao, WANG Xianling, et al. Empirical research on statistical traits of a BBS reply networks. *Complex Systems and Complexity Science*, 2010, 7(1):52-58. (in Chinese)
- [23] HEANEY M T. Multiplex networks and interest group influence reputation: an exponential random graph model. *Social Networks*, 2014, 36:66-81.
- [24] YAN L L, PENG J, TAN Y. Network dynamics: how can we find patients like us?. *Information Systems Research*, 2015, 26(3):496-512.
- [25] PAN W, SHEN C, FENG B. You get what you give: understanding reply reciprocity and social capital in online health support forums. *Journal of Health Communication*, 2017, 22(1):45-52.
- [26] BLISS C A, KLOUMANN I M, HARRIS K D, et al. Twitter reciprocal reply networks exhibit assortativity with respect to happiness. *Journal of Computational Science*, 2012, 3(5):388-397.
- [27] NEWMAN M E J, PARK J. Why social networks are different from other types of networks. *Physical Review E*, 2003, 68(3):036122-1-036122-8.
- [28] KLYMKO C, GLEICH D F, KOLDA T G. Using triangles to improve community detection in directed networks // **2014 ASE Bigdata/Socialcom/Cybersecurity Conference**. California: Stanford University, 2014:2-8.
- [29] 魏钧, 李淼淼. 团队知识转移: 多样性与网络传递性的作用. *科研管理*, 2014, 35(5):70-76.
WEI Jun, LI Miaomiao. Team knowledge transfer: the role of diversity and network transitivity. *Science Research Management*, 2014, 35(5):70-76. (in Chinese)
- [30] MCPHERSON M, SMITH-LOVIN L, COOK J M. Birds of a feather: homophily in social networks. *Annual Review of Sociology*, 2001, 27:415-444.
- [31] THELWALL M. Homophily in MySpace. *Journal of the Association for Information Science and Technology*, 2009, 60(2):219-231.
- [32] SONG X, JIANG S, YAN X, et al. Collaborative friendship networks in online healthcare communities: an exponential random graph model analysis // *Proceedings of International Conference on Smart Health*. Beijing, 2014:75-87.
- [33] GOODREAU S M, KITTS J A, MORRIS M. Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography*, 2009, 46(1):103-125.
- [34] 陶海青, 薛澜. 社会网络中的知识传递. *经济管理*, 2004(6):77-84.
TAO Haiqing, XUE Lan. The knowledge transfer in social networks. *Economic Management Journal*, 2004(6):77-84. (in Chinese)
- [35] ARGOTE L, INGRAM P, LEVINE J M, et al. Knowledge transfer in organizations: learning from the experience of others. *Organizational Behavior and Human Decision Processes*, 2000, 82(1):1-8.
- [36] LAVIE D, DRORI I. Collaborating for knowledge creation and application: the case of nanotechnology research programs. *Organization Science*, 2012, 23(3):704-724.
- [37] WU B, JIANG S, CHEN H. The impact of individual attributes on knowledge diffusion in web forums. *Quality & Quantity*, 2015, 49(6):2221-2236.
- [38] SONG X, YAN X, LI Y. Modelling liking networks in an online healthcare community: an exponential random graph model analysis approach. *Journal of Information Science*, 2014, 41(1):89-96.
- [39] SNIJDERS T A B, PATTISON P E, ROBINS G L, et al. New specifications for exponential random graph models. *Sociological Methodology*, 2006, 36(1):99-153.
- [40] ROBINS G, PATTISON P, KALISH Y, et al. An introduction to exponential random graph (p^*) models for social networks. *Social Networks*, 2007, 29(2):173-191.
- [41] WANG P, ROBINS G, PATTISON P, et al. Exponential random graph models for multilevel networks. *Social Networks*, 2013, 35(1):96-115.
- [42] PAHOR M, ŠKERLAVAJ M, DIMOVSKI V. Evidence for the network perspective on organizational learning. *Journal of the Association for Information Science and Technology*, 2008, 59(12):1985-1994.

Research on Mechanisms of User Replying Behaviors in Online Health Communities

LIU Xuan¹, WANG Linwei¹, LI Jia¹, ZHANG Pengzhu²

1 School of Business, East China University of Science and Technology, Shanghai 200237, China

2 Antai College of Economics & Management, Shanghai Jiaotong University, Shanghai 200052, China

Abstract: The persistent development of online social networks and citizens' health awareness increasingly promote the emergence

and boom of online health communities(OHCs). The online health communities provide a platform where users are able to obtain medical resources, share knowledge, experiences, and emotions with others. Recently, the OHCs is developing rapidly because of the advantages in providing abundant medical resources and health related knowledge as well as meeting users' social needs.

Among various actions proceeding in communities, posting and replying embody the establishment of social relationships, and they are fundamental for keeping boom of the communities. Thus it is worthwhile investigating the post-reply behaviors and the underline mechanisms in the communities.

We selected "Tianmijiayuan", an online diabetes health community, as our research testbed. We establish the reply networks based on the posting and replying data as well as users' personal information on this community from January, 2015 to June, 2015. The networks contain both the edges and nodes information. Then we utilize Exponential Random Graph Model (ERGM) to explore how network structure and node attributions affect the establishment of reply networks.

The results indicate the reply networks exhibit reciprocity and transitivity characteristics simultaneously. The reciprocity indicates that users tend to reply between each other, and the transitivity shows the reply networks own great intention to develop.

The influences of node attributions on reply network are much more complex. Users tend to reply those who share homophily with themselves, specifically such as the users in closed regions; users with high social capitals, such as more friends and superior activeness, would have higher possibilities to receive others' replies; those who reply others frequently and the new users are more likely to get others' replies.

This study will enrich the research paradigm of user behavior mechanisms in the communities and enrich the literatures of electronic health, social networks, online health communities and reply behaviors. In addition to the above, it will also provide enlightenment for the reply management in online communities. The results can help improve the platform design mechanism, for instance, the communities should stimulate the new users to post courageously; encourage the users who have significant influence, such as those managers, active actors, to post and reply; as well supply efficient channels to facilitate information exchange among homogenous users. Those mechanisms would further promote users' information exchange in the communities and finally determine flourish the communities.

Considering that the reply networks could be affected by other networks(such as friendship network), further research can explore the relationships among multiple networks and investigate those networks' formation mechanism. Meanwhile, replying behavior would also be affected by post's specific information, thus future research can also incorporate text mining techniques to explore the effects of post contents on replying behaviors.

Keywords: online health communities; mechanisms of reply behaviors; ERGM; network structure; node attributions

Received Date: September 10th, 2016 **Accepted Date:** December 28th, 2016

Funded Project: Supported by the National Natural Science Foundation of China(71371005,71471064,91646205), the Shanghai Pujiang Program(15PJJC019) and the Fundamental Research Funds for the Central Universities(WN1522008)

Biography: LIU Xuan, doctor in management, is an associate professor in the School of Business at East China University of Science and Technology. Her research interests include knowledge management, e-commerce and big data decision making. Her representative paper titled "Determining inventor status and its effect on knowledge diffusion: a study on nanotechnology literature from China, Russia, and India" was published in the *Journal of the American Society for Information Science and Technology*(Issue 6, 2011). E-mail:xuanliu@ecust.edu.cn

WANG Linwei is a master degree in the School of Business at East China University of Science and Technology. His research interests focus on e-health and social network analysis. E-mail:leeven92@163.com

LI Jia, doctor in management, is an associate professor in the School of Business at East China University of Science and Technology. His research interests include decision support systems, e-commerce and knowledge management. His representative paper titled "Designing an information organization structure for fuzzy tasks in online group discussion" was published in the *Journal of Management Sciences in China*(Issue 5, 2015). E-mail:jiali@ecust.edu.cn

ZHANG Pengzhu, doctor in management, is a professor in the Antai College of Economics & Management at Shanghai Jiaotong University. His research interests include e-commerce, e-government and financial information systems. His representative paper titled "G2G information sharing among government agencies" was published in the *Information & Management* (Issue 1, 2014). E-mail:pzzhang@sjtu.edu.cn □