

基于社会网络分析的网络问答社区知识传播研究

王忠义¹ 张鹤铭¹ 黄京² 李春雅³

¹(华中师范大学信息管理学院 武汉 430079)

²(武汉职业技术学院 武汉 430074)

³(南通理工学院商学院 南通 226002)

摘要:【目的】通过分析网络问答社区的社会网络结构和知识传播机制,研究节点在网络问答社区的知识传播与学习的效用,从而指导用户在网络问答社区中更高效地学习。【方法】结合社会网络分析方法和熵权法描绘意见领袖的知识能力和传播影响力,而后针对网络问答社区社会网络的小世界特征,在借鉴 Cowan 模型的基础上建立一个知识传播模型,并通过系统模拟仿真对该网络内部的知识学习效果进行实验。【结果】知识存量越少的节点在网络中的学习效率越高,该现象类似于森林的“海绵效应”;一些节点的知识存量增长较快,其知识存量的增长倍数远远大于那些当初知识存量较高的节点;随着时间的增长,网络的知识传播速率一直处于下降的趋势;知识增量与知识能力指数和传播能力指数都有很强的关系。【局限】在关于传播动力学模型的假设上,本文未考虑到网络的动态随机重连。【结论】本文较为客观地评价了网络问答社区的知识传播与学习效用的情况,并提出相应建议以提升用户在网络问答社区中的学习效果。

关键词: 社会网络分析 信息熵 知识传播 Cowan 模型 海绵效应

分类号: TP393

DOI: 10.11925/infotech.2096-3467.2018.0293

1 引言

Web2.0 环境下,互联网用户不仅仅是知识的浏览者,还是知识的生产者。这种网络模式改变了传统的知识传播形式,网络问答社区作为该知识传播形式的典型代表,它采用网络协作的方式将互联网的所有用户组成高效的社会协作网,互相解决对方提出的问题^[1]。网络问答社区是社会化网络虚拟社区的创新应用,它提供了高质量的知识分享平台,摆脱了传统问答平台的缺陷,充分体现了用户为中心和社交化理念,受到学术界和商业界的广泛关注^[2]。目前,随着海量网络信息的出现,互联网用户获取所需资源的效率引发了许多学者的思考,而用户在网络问答社区中的知识

传播和学习效果如何以及如何较好地提升学习效果等问题的探讨还较少。由此,本文以网络问答社区中的用户为研究对象,将其抽象为社会网络结构中的节点,从而分析网络问答社区的社会网络结构特征,依据网络特征构建一个知识传播模型以研究节点在网络问答社区的知识传播与学习的效用且对用户在网络问答社区中的学习效果给予评价,并提出相应建议来提升学习效果,这将对用户在网络问答社区中高效学习有较强的指导意义。

2 研究现状

网络知识问答社区作为知识交流与传播的重要平台,不仅得到用户的广泛使用,而且越来越多的学者

通讯作者: 李春雅, ORCID: 0000-0001-7155-4658, E-mail: 10880945@qq.com。

开始对其进行深入的研究。目前针对网络问答社区的相关研究主要集中于网络问答社区用户的特征分析及行为分析、知识传播模式的推断和知识传播效用的定性分析等方面。

(1) 在用户特征分析及行为分析方面, 金家华^[3]从宏观和微观两个层面探索用户在问答社区的参与行为及影响因素, 并认为基于认同的信任关系、前期知识搜寻获得的反馈、互惠规范对用户的持续知识搜寻行为有正影响; 陈娟等^[4]通过 R 语言对抓取到的社会化问答平台用户个人特征与行为的客观数据进行样本分类, 并对每一种分类的行为特征进行归纳分析; 黄梦婷等^[5]主要探讨在线问答社区用户之间的协作发起和响应、协作内容的类型以及效果的差异性并对知乎社区的协作提出一些建议; 刘佩等^[6]研究以“知乎”为代表的网络问答社区的社会网络宏观结构特征, 分析出度和入度高的用户在此社会网络中的行为特征并对网络问答社区的问题和回答内容特点进行分析; Liu 等^[7]通过训练出的回答预测模型对网络问答社区的问题回复率进行测试, 发现问题的可回答性主要取决于提问者与问题, 且如果问题带有表情符号可以增强问题被回复的可能性; Jin 等^[8]主要探讨用户为网络问答社区贡献问题回答的持续意愿的原因, 并认为其主要来源于自我展示、同行认可和社会学习的需求; Lou 等^[9]通过问卷调查的形式对主流网络问答社区的用户进行样本采集, 并分析出激励用户为网络问答社区贡献回答的因素主要是: 名誉奖励系统、学习知识的效用和乐意帮助

的心理。

(2) 针对知识传播模式的推断和知识传播效用的定性分析方面, 左遥等^[10]通过归纳出基本的知识传播方式从而提出一种社会化问答网站知识传播网络推断方法, 分别提取节点类型、节点时间关系及候选传播者排序特征推断关注关系, 采用启发式方法推断另外三种关系, 生成知识传播网络; 宁菁菁^[11]通过“弱关系理论”分析知乎的知识传播机制并对其知识传播特征和效果进行定性的分析。

综上, 针对网络问答社区知识传播的特征模型构建、分析和学习效用评价的相关研究还不完善, 尤其对于其知识传播的模型构建和定量分析的研究较少。因此, 本文通过借鉴 Cowan 模型^[12]深入剖析网络问答社区知识传播机制, 并建立新模型完成仿真模拟知识学习效果, 从而对网络问答社区的学习效用做出评价并给出相应建议, 这有利于指导用户在网络问答社区中更高效地获取知识。

3 网络问答社区知识传播研究框架

为了实现对网络问答社区的传播模型构建、分析和学习效用评价, 本文大致沿袭了数据采集、数据预处理、数据使用和数据分析的一般范式设计研究框架, 如图 1 所示。

- (1) 数据收集: 获取用户的相关数据;
- (2) 社会网络构建: 根据收集数据中用户的相互关注关系构建社会关系网络;
- (3) 社会网络宏观结构分析: 从社会网络的整体

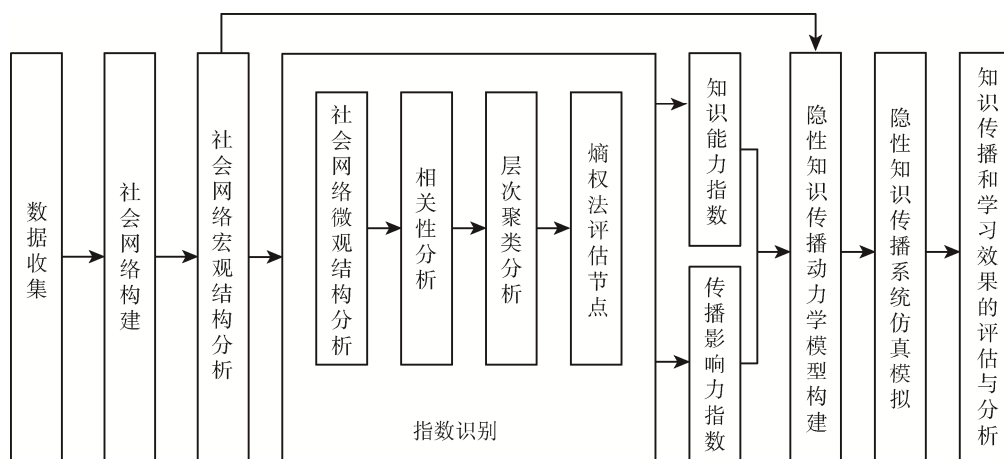


图 1 网络问答社区知识传播研究框架

出发,使用 R 语言实现社会网络特性分析;

(4) 指数识别:使用 R 语言对社会网络的节点进行相关参数分析,在已获得的用户特征数据和社会网络微观结构分析的数据中,进行相关性分析和层次聚类分析,将数据聚类为知识能力和传播影响力。在此基础上,结合熵权法得出用户的知识能力指数和传播影响力指数;

(5) 隐性知识传播动力学模型构建:结合之前的社会网络特征分析,针对其特征在改进 Cowan 模型的基础上构建知识传播动力学模型;

(6) 隐性知识传播系统仿真模拟:将已处理的数据结合动力学模型,对用户知识传播与学习的行为通过 Matlab^[13]进行系统仿真模拟;

(7) 知识传播和学习效果的评估与分析:通过分析系统的仿真模拟结果,得出相关实验结论并对网络问答社区的知识学习效果做出评价和提出相关建议。

4 数据收集以及网络构建

4.1 数据收集

“知乎”作为目前国内最受欢迎的知识型网络问答社区^[6],采用精英与专家相结合的路线,其高质量的问答、理性友好的讨论氛围等方式使得其在蓬勃发展的网络问答社区中脱颖而出。由于最大程度调动了用户甄别问题的质量,“知乎”构建了高质量问答社区^[1]。为此,本文以“知乎”社区作为研究对象,探索网络问答社区知识传播和学习效果。

本文从“知乎”共爬取420 981个用户数据,爬虫的截止日期为2017年2月。在得到用户的数据后,针对获得的数据粗略筛选出一定数量的意见领袖。所谓网络意见领袖,是指以互联网为平台,针对社会热点问题发表自己的言论和见解,并得到网民普遍认可或高度认同,从而引导网络舆论走向,对网民意见产生重大影响的网络信息发布者^[14]。基于以上对意见领袖标准的描述,借鉴美国加利福尼亚大学圣地亚哥分校的物理学家乔治·赫希在2005年提出的H指数(H-index)对知乎用户进行科学筛选。H指数是一种科学评价计量,该指数既能反映学术论文影响力大小,又能同时反映其重要论文产出数量的多少^[15]。本文基于H指数的核心思想,从两个维度衡量意见领袖:被关注人数以及获得赞同数量,将以上两个数量上均大于100 000

的用户筛选出来作为意见领袖,总共得出80位“知乎”意见领袖,如表1所示。

4.2 社会网络构建

将“知乎”用户抽象为社会网络中的节点,基于用户关注与被关注关系构建关注关系矩阵,关注某位用户则表示为1,不关注则表示为0。由于数据庞大,这里仅展示80位意见领袖的前20位的有向关注关系矩阵,如表2所示。

将关注关系矩阵导入Rstudio^[16],并在R语言中借助iGraph包建立有向图函数。在函数中,各个节点是被抽象化的用户,关注关系是依据矩阵每行中的0和1确定两个节点之间是否连线,并且依据“知乎”情境,将函数direct值设置为T来建立有向关注关系。在建立有向图的同时,为了能够更加形象化地描述节点的重要性程度,对于节点的大小,依据节点的中间中心度属性来调节。中间中心度大于100的节点,节点size属性设置为16;小于等于30的节点, size属性设置为6;其余节点的size属性都设置为默认的10。这样即可绘制出意见领袖的社会网络关系的有向图,如图2所示。

4.3 社会网络宏观结构分析

经过R语言中graph.Density, mean_distance和transitivity函数的计算,分别得到该社会网络的网络密度:0.3753165;网络平均路径长度:1.649209;网络聚集系数:0.6202949。通过这些数据可以认为,该网络符合小世界网络高聚集系数、低平均路径长度的特征^[17],网络密度比较大,成员之间的关系比较紧密。此外,从社会网络关系图中可以明显看出核心-边缘结构,核心节点对于该社会网络有较强的影响力。意见领袖的关注关系网络的小世界性质为隐性知识传播动力学模型的构建奠定了理论基础。

5 指数识别

5.1 社会网络微观结构分析

本文从“知乎”共爬取19个用户字段,如表3所示。去除不是主键的字段:address, education, company, job, headline, 删去重复标记的主键:user_id, index_url, 再删除对评价用户不重要或冗余的字段:sex, weibo, token, 最终抽取其中9个字段刻画用户的特征,如表4所示。

表 1 80 位意见领袖, 按赞同数量(agree)排名

序号	username	agree	followers	序号	username	agree	followers
01	肥肥猫	1 218 509	353 535	41	warfalcon	246 569	478 527
02	朱炫	1 128 626	579 459	42	孙志超	239 546	185 374
03	ze ran	1 029 394	207 678	43	马力	235 034	204 201
04	vczh	983 884	467 598	44	藥師	233 497	112 765
05	寺主人	940 156	454 217	45	闻佳	227 652	294 435
06	Hannibal Lecter	841 383	222 944	46	蔓玫	219 097	126 074
07	yolfilm	835 981	732 463	47	高科	218 227	505 662
08	菠菜	685 089	340 044	48	汪惟	212 827	306 975
09	孟德尔	646 872	195 376	49	pansz	208 025	212 482
10	Kaiser	635 745	277 571	50	Lightwing	202 779	132 399
11	一笑风云过	628 206	181 991	51	带三个表	199 884	312 977
12	银教授	603 600	270 015	52	沃金	196 759	118 629
13	曾加	581 484	204 864	53	李楠	190 009	415 213
14	谢熊猫君	575 400	389 782	54	张亮	187 148	697 974
15	Justin Lee	526 041	160 485	55	何明科	181 619	129 164
16	君临	483 966	182 307	56	ALEX YA	181 589	104 854
17	windleavez	463 623	106 680	57	负二	180 996	398 218
18	护耳大脸	453 444	120 567	58	小岩井	177 677	114 498
19	李松蔚	432 592	305 226	59	李暘	171 104	128 141
20	苏菲	418 907	342 303	60	葛巾	168 827	580 650
21	安雅	392 539	147 147	61	maggie	168 648	552 459
22	王路	377 760	107 665	62	Sophia	165 521	101 880
23	梁边妖	361 965	540 804	63	不鳥萬如一	160 960	397 024
24	倪一宁	360 068	153 023	64	欲三更	151 195	158 159
25	李森	347 455	623 385	65	喻忘忧	146 550	106 649
26	极乐	344 128	114 508	66	楚沐风	146 530	119 321
27	蒋校长	341 507	152 871	67	Raymond Wang	144 990	484 481
28	动机在杭州	310 503	452 619	68	殷守甫	142 391	121 032
29	刘鹏程 Sai.L	309 067	165 133	69	徐强	141 421	148 944
30	刘念	302 048	267 137	70	纽约老李校长	137 423	136 041
31	猪小宝	296 676	134 047	71	东东枪	135 709	198 629
32	徐湘楠	295 563	106 909	72	唐僧同志	133 168	126 723
33	陈章鱼	291 457	516 409	73	楠谷	131 511	129 417
34	cCOMMANDO	289 557	355 508	74	覃超	118 423	123 436
35	顾扯淡	283 796	236 151	75	周晓农	118 290	481 778
36	夏吉吉	273 456	144 224	76	涛吴	116 545	182 749
37	命硬的 eno	270 775	162 779	77	谭蔓茹	114 534	152 995
38	David Rand	265 537	205 266	78	Lydia	112 501	123 632
39	雷幺幺	262 207	115 092	79	李开复	108 208	981 917
40	汗青	257 863	111 979	80	夏昊 BFA	101 281	197 867

表 2 前 20 位意见领袖的关注关系矩阵

序号	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20
01	0	1	1	0	0	0	0	1	0	1	1	0	0	0	1	1	0	1	0	0
02	1	0	0	0	0	0	0	1	0	1	1	1	0	1	0	0	1	0	1	0
03	1	0	0	0	0	0	0	1	0	0	1	0	1	1	1	1	0	0	1	0
04	0	0	0	0	1	0	1	1	1	0	1	0	0	0	0	0	1	0	0	1
05	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
06	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	0	0	1	0
07	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	1	0	0	1	0
08	1	1	1	1	0	1	1	0	0	0	1	0	1	1	1	0	0	1	1	0
09	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0
10	1	1	0	1	0	1	0	0	0	0	1	0	1	0	0	1	1	0	0	0
11	1	1	1	0	1	0	1	1	0	0	0	0	1	1	1	0	1	0	1	1
12	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0
13	1	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0
14	0	1	1	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	1	1
15	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	0	1	0
16	1	1	0	0	1	0	1	1	0	1	0	0	0	1	1	0	0	0	1	1
17	0	1	0	0	0	0	1	0	1	1	0	0	1	0	0	0	0	0	0	0
18	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	1	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	1
20	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0

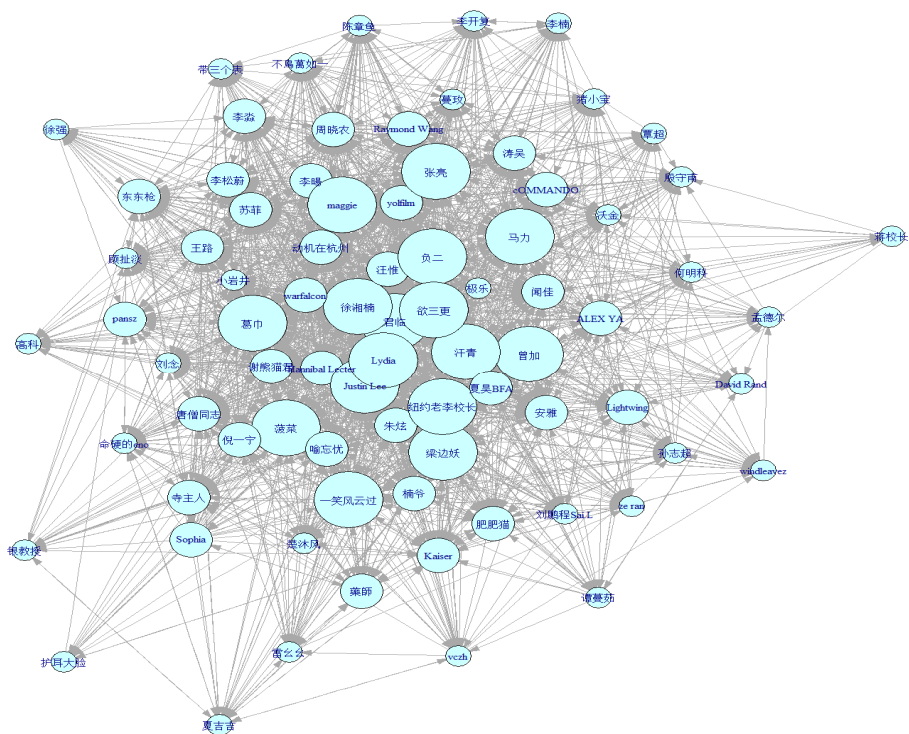


图 2 “知乎”社区 80 位意见领袖的关系网络

表 3 从“知乎”获取的 19 个字段名称

字段名									
username	company	user_id	article	education	question	thanked	topic	followers	weibo
address	job	answer	index_url	headline	agree	following	token	sex	

表 4 前 10 位意见领袖特征描述

username	answer	question	article	agree	thanked	followers	following	topic
肥肥猫	397	89	12	1 218 509	304 153	353 535	257	11
朱炫	196	4	47	1 128 626	245 011	579 459	204	37
ze ran	2 519	0	6	1 029 394	198 023	207 678	703	137
vczh	13 688	459	46	983 884	130 151	467 598	1 841	30
寺主人	149	41	156	940 156	192 639	454 217	714	58
Hannibal Lecter	972	32	149	841 383	188 170	222 944	16 884	292
yolfilm	1 509	106	2	835 981	198 641	732 463	226	134
菠菜	1 295	19	37	685 089	190 034	340 044	536	49
孟德尔	5 096	861	47	646 872	141 033	195 376	122	8

为获取“知乎”中意见领袖的影响力特征，获得用户中间中心度 (betweeness) 数据后，进行社会网络微观结构分析，以获取意见领袖的出度 (degree.out)，入度 (degree.in)，点度中心度 (degree.

total)，接近中心度 (closeness)，特征向量中心度 (evcent)，pagerank 指数，结构洞指数 (shole) 等，如表 5 所示。由此，可以更加全面地描述意见领袖的影响力。

表 5 前 10 位意见领袖节点影响力分析

username	betweeness	degree.out	degree.in	degree. total	closeness	evcent	pagerank	shole
肥肥猫	63.96	26	35	61	0.0076	0.1075	0.0137	36
朱炫	95.05	35	47	82	0.0081	0.1442	0.0181	52
ze ran	11.42	14	20	34	0.0069	0.0648	0.0076	21
vczh	15.67	14	20	34	0.0069	0.0589	0.0075	29
寺主人	34.57	8	35	43	0.0065	0.0737	0.0135	38
Hannibal Lecter	72.48	48	22	70	0.0091	0.1187	0.0099	51
yolfilm	94.76	29	48	77	0.0077	0.1336	0.0218	48
菠菜	120.51	43	44	87	0.0087	0.1524	0.0166	53
孟德尔	3.47	6	20	26	0.0060	0.0484	0.0074	23

汇总发现以上数据共有 16 维。考虑到维度过高，不利于数据分析，将通过一些数据降维的方法概括用户的属性，方便进一步分析用户特征。

5.2 相关性分析

首先，通过相关性分析合并关系比较紧密的数据。为此，借助 R 语言中 cordata 函数对 16 维数据进行相关分析，相关性系数如表 6 所示。从相关分析的

结果可以发现赞同数与感谢数有很强关系；中间中心度、出度、入度、点度中心度、接近中心度、特征向量中心度、pagerank、结构洞指数之间有很强的相关性(见表 6 中加粗数据)。

为更加显示地刻画参数之间的相关性程度，借助 R 语言 car 包中的 scatterplotMatrix 函数分别绘制出中间中心度、出度、入度、点度中心度、接近中心度、

特征向量中心度、pagerank、结构洞指数之间的相关性,如图3所示。根据表6和图3,选择筛选掉相关性较大的维度和冗余的数据,保留其中之一即可。于是,在16个维度的基础上,本文采用赞同数量(agree)代替感谢数量(thanked),出度(degree.out)代替接近中心度

(closeness),入度(degree.in)代替 pagerank 指数,点度中心度(degree.total)代替特征向量中心度(evcent),而关注数量(following)和被关注数量(followed)可以通过其他影响力指数体现,在这里可以直接删去,最后保留10个维度。

表6 80位意见领袖16维数据的相关性

	answer	question	article	agree	thanked	followers	following	topic	betweeness	degree.out	degree.in	degree.total	closeness	evcent	pagerank	shole
answer	1.00	0.39	0.12	0.29	0.12	0.17	0.09	-0.01	-0.09	-0.17	-0.06	-0.15	-0.15	-0.16	-0.05	-0.10
question	0.39	1.00	0.29	-0.02	-0.06	0.29	0.05	0.15	0.07	-0.08	0.11	-0.01	-0.09	-0.02	0.21	0.02
article	0.12	0.29	1.00	-0.07	-0.14	-0.03	0.04	0.09	0.02	-0.04	0.04	-0.01	-0.05	-0.01	0.11	-0.02
agree	0.29	-0.02	-0.07	1.00	0.92	0.20	0.14	-0.05	0.08	-0.07	0.07	-0.02	-0.06	-0.03	0.00	-0.03
thanked	0.12	-0.06	-0.14	0.92	1.00	0.24	0.13	0.01	0.16	0.03	0.15	0.10	0.04	0.09	0.08	0.07
followers	0.17	0.29	-0.03	0.20	0.24	1.00	-0.02	0.09	0.17	-0.12	0.52	0.17	-0.16	0.18	0.56	0.14
following	0.09	0.05	0.04	0.14	0.13	-0.02	1.00	0.25	0.19	0.22	-0.01	0.16	0.23	0.13	0.01	0.20
topic	-0.01	0.15	0.09	-0.05	0.01	0.09	0.25	1.00	0.13	0.25	0.05	0.21	0.24	0.20	0.03	0.23
betweeness	-0.09	0.07	0.02	0.08	0.16	0.17	0.19	0.13	1.00	0.76	0.66	0.88	0.73	0.85	0.65	0.81
degree.out	-0.17	-0.08	-0.04	-0.07	0.03	-0.12	0.22	0.25	0.76	1.00	0.30	0.88	0.99	0.85	0.25	0.90
degree.in	-0.06	0.11	0.04	0.07	0.15	0.52	-0.01	0.05	0.66	0.30	1.00	0.71	0.24	0.74	0.94	0.57
degree.total	-0.15	-0.01	-0.01	-0.02	0.10	0.17	0.16	0.21	0.88	0.88	0.71	1.00	0.85	0.99	0.65	0.94
closeness	-0.15	-0.09	-0.05	-0.06	0.04	-0.16	0.23	0.24	0.73	0.99	0.24	0.85	1.00	0.81	0.20	0.88
evcent	-0.16	-0.02	-0.01	-0.03	0.09	0.18	0.13	0.20	0.85	0.85	0.74	0.99	0.81	1.00	0.67	0.91
pagerank	-0.05	0.21	0.11	0.00	0.08	0.56	0.01	0.03	0.65	0.25	0.94	0.65	0.20	0.67	1.00	0.53
shole	-0.10	0.02	-0.02	-0.03	0.07	0.14	0.20	0.23	0.81	0.90	0.57	0.94	0.88	0.91	0.53	1.00

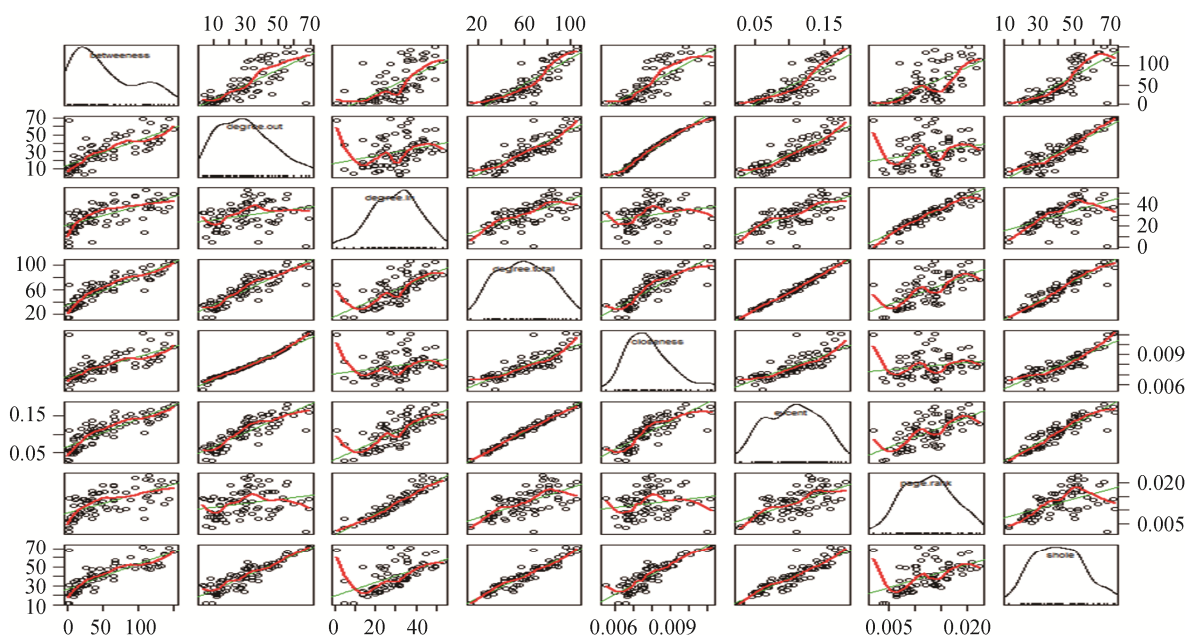


图3 节点影响力分析的8维数据相关性

5.3 层次聚类分析

为深入揭示 10 个维数据的共性特征, 实现数据的进一步降维, 通过层次聚类法对其进行聚类分析。在层次聚类之前, 借助 R 语言中线性归一化函数^[18](见公

式(1))对数据进行归一化处理以统一量纲; 再利用 hclust 函数对数据实现欧氏距离聚类(见图 4)。

$$x' = \frac{\max(x) - x}{\max(x) - \min(x)} \tag{1}$$

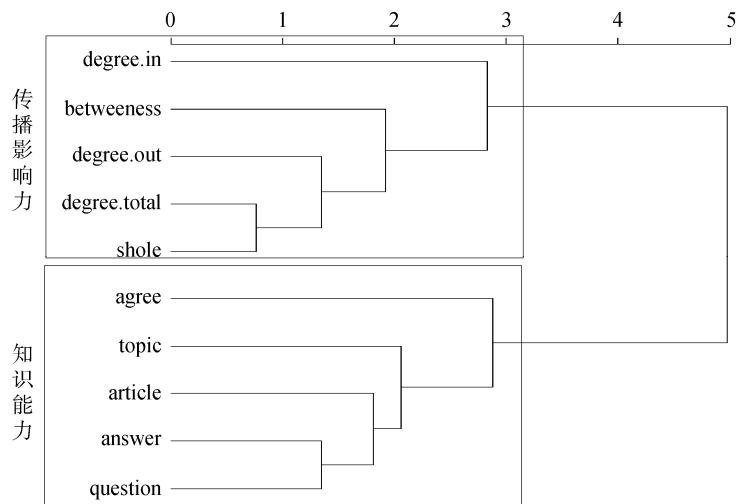


图 4 10 个维度数据层次聚类结果

从图 4 可以看出, 分析节点影响力的 5 个维度数据(即, degree.in、betweenness、degree.out、degree.total、shole)之间有很强的相关性。除此之外, 其他 5 个用户属性维度(即, agree、topic、article、answer、question)之间的聚类效果也很明显。基于以上聚类的结果, 本文将分析用户节点影响力的 5 个维度数据归类为传播影响力, 将描述用户特征的这 5 个维度归类为知识能力。

5.4 利用熵权法评估节点

为了能够更加客观考量每个因素的权重, 本文引入熵权法确定权重。信息熵是消除不确定性所需信息量的度量, 即未知事件可能含有的信息量^[19]。应用信息熵可以度量获取数据所提供的信息量, 并依据此确定各指标的权重^[20]。信息熵的计算方法^[20]如公式(2)所示。权重计算方法^[19]如公式(3)所示。

$$Entropy(j) = \ln m \times \sum_{i=1}^m \left(\frac{d_{i,j}}{\sum_{i=1}^m d_{i,j}} \times \ln \left(\frac{d_{i,j}}{\sum_{i=1}^m d_{i,j}} \right) \right) \tag{2}$$

$$j = 1, 2, \dots, n, d_{i,j} \text{ 有 } m \text{ 行 } n \text{ 列}, m = 80, n = 10$$

$$Weight(j) = \frac{1 - Entropy(j)}{n - \sum_{j=1}^n Entropy(j)} \tag{3}$$

其中, j 表示每一个因素的信息熵; i 表示每个数据维度下 80 位意见领袖的相关因素数据。借助 Matlab 可以计算出每个因素的信息熵和权重, 如表 7 所示。

表 7 信息熵数值和熵权数值

因素	信息熵	权重	因素	信息熵	权重
answer	0.8081	0.1949	betweenness	0.9170	0.3901
question	0.6867	0.3182	degree.out	0.9527	0.2223
article	0.7665	0.2371	degree.in	0.9798	0.0949
agree	0.8935	0.1082	degree.total	0.9680	0.1506
topic	0.8607	0.1415	shole	0.9698	0.1421

在得到这 10 维数据在各自聚类中的权重之后, 对之前归一化后的数据使用加权求和的方式最终得到这 80 位意见领袖的知识能力指数和传播影响力指数, 如表 8 所示。

为了对知识能力和传播影响力做出评价, 绘制图 5 所示的分布图。可以发现传播影响力指数分布不是很集中, 每个分数段的人群数量比较平均。知识能力指数的分布是一个左偏态的分布, 说明知识能力的分布不平均, 个体之间的差异较大, 这为用户在这样的社区内学习提供了广阔的空间。

表 8 知识能力指数和传播影响力指数

用户	知识能力	传播影响力	用户	知识能力	传播影响力	用户	知识能力	传播影响力
肥肥猫	23.77	45.10	动机在杭州	9.24	64.57	何明科	9.86	20.25
朱炫	21.29	67.28	刘鹏程 Sai.L	26.33	20.75	ALEX YA	2.99	44.67
ze ran	28.03	14.39	刘念	6.18	33.14	负二	19.19	98.45
vczh	70.35	17.63	猪小宝	8.56	20.75	小岩井	2.37	43.71
寺主人	24.75	27.61	徐湘楠	23.89	91.96	李暘	16.75	40.58
Hannibal Lecter	34.53	58.36	陈章鱼	15.28	20.29	葛巾	0.74	72.70
yolfilm	25.20	63.31	cOMMANDO	11.21	41.45	maggie	6.02	80.88
菠菜	16.77	77.90	顾扯淡	9.85	24.32	Sophia	1.39	31.76
孟德尔	54.34	8.34	夏吉吉	2.64	10.67	不鳥萬如一	95.00	26.46
Kaiser	11.74	44.06	命硬的 eno	2.79	32.39	欲三更	12.96	86.40
一笑风云过	16.73	63.06	David Rand	12.15	13.62	喻忘忧	1.95	63.49
银教授	13.41	8.72	雷么么	3.66	14.68	楚沐风	2.91	23.91
曾加	12.03	71.10	汗青	12.74	80.08	Raymond Wang	8.00	53.73
谢熊猫君	13.53	57.75	warfalcon	29.78	73.38	殷守甫	1.56	18.04
Justin Lee	7.83	100.00	孙志超	26.48	19.28	徐强	8.18	8.09
君临	8.58	78.61	马力	60.21	70.62	纽约老李校长	10.31	90.03
windleavez	15.21	11.69	藥師	13.46	32.31	东东枪	11.01	35.73
护耳大脸	10.97	0.00	闻佳	4.14	38.34	唐僧同志	2.94	46.39
李松蔚	10.96	44.45	蔓玫	5.04	29.77	楠爷	3.87	43.25
苏菲	12.69	45.34	高科	6.41	10.27	覃超	9.16	14.32
安雅	9.81	37.48	汪惟	10.26	64.13	周晓农	8.97	47.54
王路	22.39	38.22	pansz	13.04	34.85	涛吴	10.29	37.20
梁边妖	10.26	79.40	Lightwing	3.68	31.17	谭蔓茹	2.61	13.39
倪一宁	3.76	51.21	带三个表	9.27	19.16	Lydia	2.63	89.12
李淼	21.23	39.31	沃金	2.79	29.81	李开复	31.16	13.02
极乐	10.12	45.18	李楠	23.13	16.28	夏昊 BFA	35.45	62.92
蒋校长	8.92	1.13	张亮	72.12	70.08			

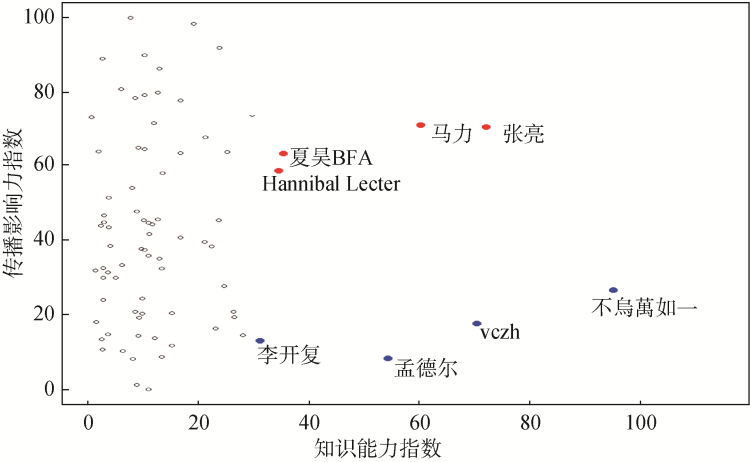


图 5 80 位意见领袖权威指数评估

6 隐性知识传播动力学分析

6.1 隐性知识传播动力学模型构建

本文在社会网络宏观结构分析中指出网络问答社区社会网络符合小世界模型高聚集系数、低平均路径长度的特征,针对此种特性,从香农模型^[21]、SIS 传染病模型^[22]和 Cowan 知识传播模型中,笔者选取 Cowan 知识传播模型进行模型构建。

选择 Cowan 模型的主要原因在于需要研究个体节点的知识学习效果,而香农模型忽视了节点之间客观存在的影响力^[23],SIS 传染病模型只能粗略描述节点是否拥有某一知识,对于揭示节点知识存量的增长率和增长量能力较弱,而 Cowan 模型很好地考虑到节点之间的隐性知识传播作用关系、能更好地反映知识网络微观变化等方面的原因。但 Cowan 模型没有考虑用户的知识沉淀和隐性知识表出化需要消耗一定的时间,因此本文增加了假设(1)。此外,考虑到要将模型置于网络问答社区的知识传播过程中,本文对 Cowan 模型又增加了假设(2)、假设(3)和假设(4)。

(1) 考虑到知识沉淀和隐性知识外化的时间消耗

过程,设定学到的知识要经过两个单位时间的延迟才可以传播;

(2) 考虑到意见领袖的基础知识存量差异较小,认为节点在知识前沿领域的知识存量就是节点的知识能力指数;

(3) 考虑到由于知识势差,知识会自动从知识存量高的节点传播至知识存量较低的节点^[24],认为节点之间知识存量存在差异就存在知识传播;

(4) 考虑到系统仿真时间长度较短的原因,并没有设置知识存量的上限。

隐性知识传播动力学模型如图 6 所示,可以看出在 $t=1$ 时,节点 j 通过自学习,将自身的知识存量 $k_j(t)$ 增加到 $n_j(t+1)$ 。由于隐性知识显化需要一定时间,所以本文设定延迟时间为两个单位时间,因此在这次学习过程中,其他节点如节点 i 将自己外化的知识 $k_i(t)$ 传授给节点 j ,使得节点 j 的知识存量增加到 $k_j(t+1)$ 。在 $t=2,3,\dots$ 时刻,节点 j 同样通过自学习,将自身的知识存量 $k_j(t+1)$ 增加到 $n_j(t+2)$ 。在此次学习过程中,其他节点如节点 i 将自己外化的知识 $k_i(t)$ 传授给节点 j ,使得节点 j 的知识存量从 $n_j(t+2)$ 增加到 $k_j(t+2)$ 。

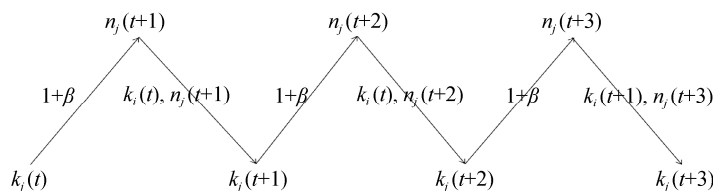


图 6 1 阶时间延迟的隐性知识传播模型示意图

在上述的知识转化过程中,必须存在节点 i 对节点 j 的知识势差,籍此,知识才能发生传递。本文将上述过程在 Cowan 模型的基础上进行改进作为新的传播动力学模型,如以下方程组所示。

$$\begin{cases} n_j(t+2) = k_j(t+1)(1+\beta) \end{cases} \quad (4)$$

$$\begin{cases} k_j(t+2) = n_j(t+2)(1+g(k_i(t), n_j(t+2))) \times P_{j,i} \\ P \text{ 为邻接矩阵} \end{cases} \quad (5)$$

$$\begin{cases} g(k_i(t), n_j(t+2)) = \max \{0; r_{j,i}(1-r_{j,i})\} \end{cases} \quad (6)$$

$$\begin{cases} r_{j,i} = \frac{k_i(t)}{n_j(t+2)} \end{cases} \quad (7)$$

$$\begin{cases} t = 2, 3, \dots \end{cases} \quad (8)$$

上述方程组中,式(4)表示节点 j 通过自学习,将自身知识从 $k_j(t+1)$ 提升至 $n_j(t+2)$,其中自学习增长率为 β ;式(5)–式(7)联立,可以计算出节点 j 在一个单位时间后经由函数 g 得到的传播率而最终得到的知识存量,这里考虑到两个单位时间的延迟,将传输节点两个单位时间之前未自学习的知识存量 $k_i(t)$ 与节点 j 此时刻的知识存量的比率作为判定标准,并依据最大值函数从而判定两个节点之间是否存在知识势差。最后依据两个节点之间的邻接矩阵判断两个节点之间是否存在知识传播关系,从而得到最终节点 j 在经过一个单位时间后的知识存量。还需要说明的是,虽然式(8)

是从 $t=2$ 开始计算的,但是在现实情境中, $t=1$ 是具有实际意义的,在算法实施的时候 $t=1$ 需要另外考虑,从 $t=2$ 开始的数据才具备仿真模拟效果,所以在最终的绘图中(见图 7 至图 9),都是从 $t=2$ 时刻开始。

6.2 隐性知识传播系统仿真模拟

为了能够对本网络用户知识传播与知识学习效用进行更深入的分析,通过计算机系统仿真模拟现实中知识传播与学习的形势演化及其发展过程^[25],因此设计了计算机系统仿真模拟的算法,如算法 1 和算法 2 所示。

初始参量: K 是知识存量矩阵,在 $t=1$ 时赋初值; N 是自学习后知识存量矩阵; $\beta=0.0005$ 是节点自学习知识增长率(β 值经过多次调试,这里选择一个较符合现实情境的值); T 为仿真模拟次数; G 是每个时期节点知识存量增长率; P 是邻接矩阵(即关注关系矩阵)。

算法 1: 隐性知识传播算法

```

1:  $t=1$ 
2: for  $t \leq T$  do
3:    $N^{(t+1)} = K^{(t)} \times (1 + \beta)$  //节点通过自学习增长知识存量
4:   for 每个  $k_j^{(t)} \in K^{(t)}$  do //针对  $t$  时刻节点  $j$  的知识存量变化
5:     for 每个  $p_{j,i} \in P_j$  do
6:       if  $t=1$ 
7:         call  $g(k_i^{(t)}, n_j^{(t+1)})$  //  $g$  入  $g$  函数
8:        $G_j^{(t)} = G_j^{(t)} + g(k_i^{(t)}, n_j^{(t+1)}) \times p_{j,i}$ 
9:     else
10:      call  $g(k_i^{(t-1)}, n_j^{(t+1)})$ 

```

```

11:     $G_j^{(t)} = G_j^{(t)} + g(k_i^{(t-1)}, n_j^{(t+1)}) \times p_{j,i}$ 
12:  end if
13: end for
14:  $k_j^{(t+1)} = n_j^{(t+1)} \times (1 + G_j^{(t)})$  //节点  $j$  在  $t$  时刻的知识存量增长
15: end for
16:  $t=t+1$ 
17: end for

```

算法 2: g 函数生成过程

```

1:  $k'$  是算法 1 中  $g$  函数的第一个传递参数,  $n'$  是第二个传递参数
2:  $ratio = k' / n'$ 
3: if  $ratio \times (1 - ratio) < 0$ 
4:    $g(k', n') = 0$ 
5: else
6:    $g(k', n') = ratio \times (1 - ratio)$ 
7: end if

```

将系统模型的方程转化成 Matlab 代码,通过 Matlab 系统仿真来模拟意见领袖知识学习效果,结果如图 7 至图 9 所示。每个子图中所展现的数据的计算方式如下:

- ①整个网络每个时刻净学习增量 $\frac{1}{80} \sum_{i=1}^{80} [k_i(t+1) - n_i(t+1)]$ ($t=1, 2, \dots$);
- ②整个网络每个时刻知识传播速率 $\frac{1}{80} \sum_{i=1}^{80} G_i(t)$ ($t=1, 2, \dots$);
- ③每个节点知识学习速率的平均值 $\frac{1}{t} \sum_{i=1}^t G_i(t)$ ($i=1, 2, \dots$);
- ④每个节点知识存量的增长倍数 $\frac{k_i(t+1)}{k_i(1)}$ ($i=1, 2, \dots$)。

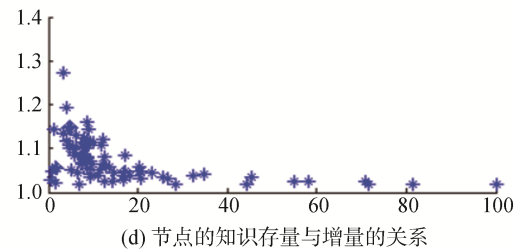
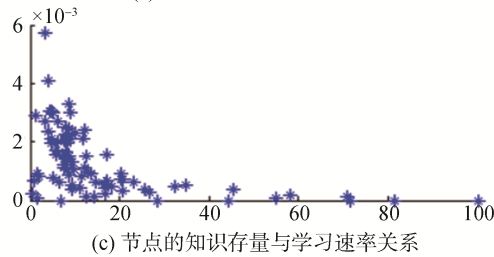
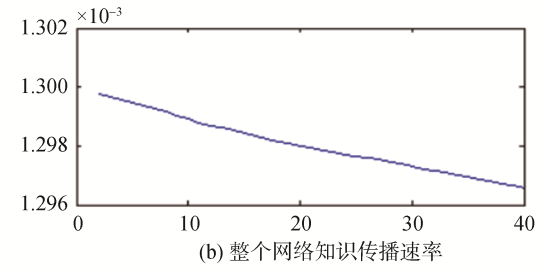
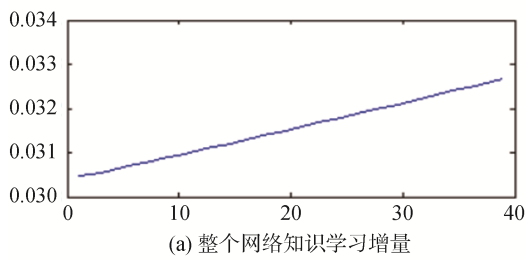
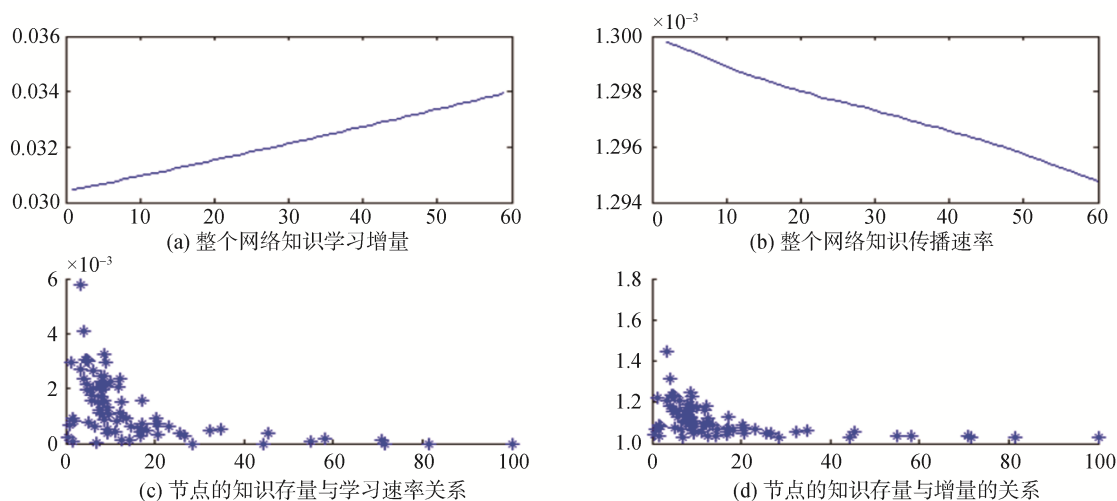
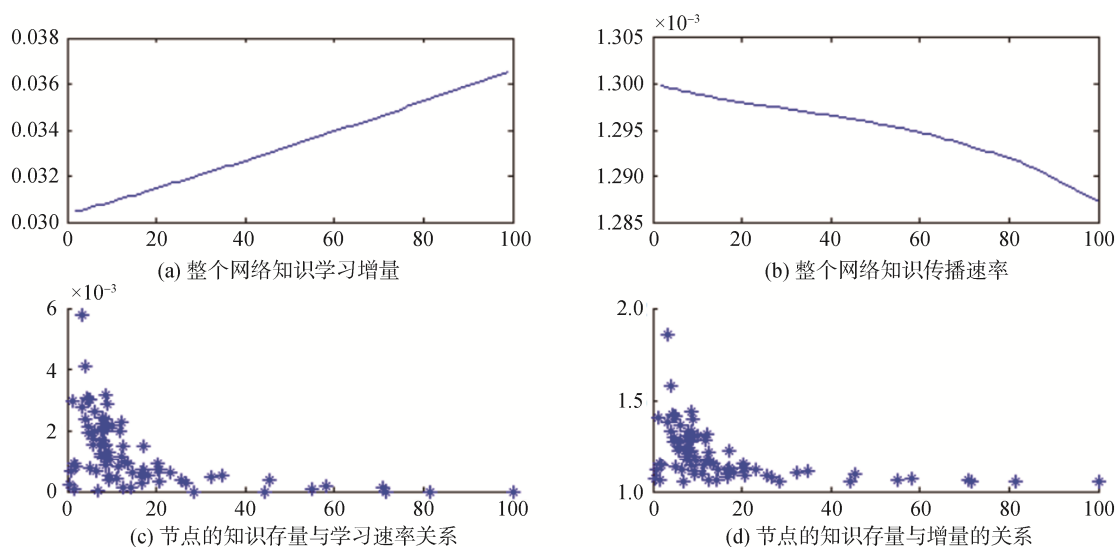


图 7 $t=40$ 时系统仿真模拟结果

图 8 $t=60$ 时系统仿真模拟结果图 9 $t=100$ 时系统仿真模拟结果

6.3 隐性知识传播系统仿真模拟结果定性分析

(1) 节点知识存量越低, 节点的学习速率越高, 同时节点的知识增量也越高。这符合人们对于学习效果的基本认知, 即当知识存量较少的人在同比其知识存量高许多的人群处在一起, 可以显著地提升他们的知识存量。但是在知识存量增长倍数方面, 随着时间的增长, 增长最快的一部分群体是起始知识存量在 10 左右的群体, 这意味着进入这样一个群体学习, 低知识存量节点的学习效果要明显优胜于高知识存量的节点。但是, 这是基于一定前提的, 知识存量过低意味着学习能力不强, 过早进入这种层次的网络, 学习效果也不会很凸显。

(2) 有些节点是具备后发优势的。他们由于知识存量较少便在网络中潜心学习, 在经过一段单位时间的学习后, 其知识存量的增长倍数远远大于那些当初知识存量较高的节点。这说明, 一些节点在该网络中的知识话语权可能还会逐渐改变, 一些节点可能会成为该网络中的新兴知识权力节点, 而一些节点可能会因为骄傲或偏见而逐渐被边缘化。

(3) 在观测 $t=100$ 时整个网络的知识传播速率可以发现, 随着时间的增长, 网络的知识传播速率一直处于下降的趋势。这说明, 长期处于某一固定群体当中, 用户的知识水平会趋于比较平稳的状态, 中间层次的用户之间的差异会相对减少。而中间层次用户差

异的减少必然会导致用户对知识学习的“免疫”，从而对网络中的知识产生厌倦。这启示用户要不断更新自己的关注对象，及时关注网络中的新兴知识权力节点，从而使得自身的学习效果达到更好的状态。

综上，以上现象与森林的“海绵效应”类似，在节点自身的知识存量低于外界时，会主动吸收知识，而且节点知识存量越少其吸收速率越高；反之，节点自身的知识存量高于外界时，会选择放水。但是，当一个节点的知识存量过低时，节点自身的学习能力极其微弱，就像一块干燥坚硬的海绵，无论其吸收速率还是吸收增量都是极低的。而如果长期处于同一种环境下，用户间的知识水平会逐渐趋于平均，使得整个网络的知识传播速率下降。本文将这种现象称为知识的“海绵效应”。

6.4 隐性知识传播系统仿真模拟结果定量分析

出于计算精度的考虑，这里选取 $t=40$ 时刻，知识增量与学习速率同二维评价指标和其下 10 维数据之间的回归分析，如表 9 所示。通过 R 语言中的多元线性回归函数 lm，前后进行两次多元线性回归后分析得出的结论是：知识增量与知识能力指数和传播能力指数都有很强的关系。其中，提问数量(question)、赞同数(agree)和出度(degree.out)对于知识增量的影响力最大。

表 9 多元线性回归参数

	回归系数	t 值	p 值
知识能力指数	0.8020	3.090	0.00278
传播影响力指数	1.2039	7.106	5.22e-10
answer	0.002676	1.002	0.31996
question	0.05065	2.654	0.00984
article	0.03714	1.567	0.12166
agree	0.00003078	1.808	0.07494
topic	-0.05145	-1.257	0.21310
betweeness	0.01136	0.056	0.95585
degree.out	1.521	1.888	0.06313
degree.in	-0.6753	-0.981	0.33001
degree.total	1.1364	0.616	0.54021
shole	1.088	1.149	0.25456

基于多元线性回归的初步结论，认为网络问答社区内，用户通过多提问题、多关注高质量用户和浏览

他人的动态、多回答高质量问题，可以明显地提升自身的知识量。

提问是一种积极与其他用户交互的过程，在“知乎”这样的问答社区中，提问是用户获取知识的最重要的方式之一。通过提问，用户会自发整理自身的知识，这本身就是对自身知识的一种回顾，不仅可以温习以往知识，还可以锻炼用户的语言表达能力和促进用户的思维发展。而且，这种获取知识的方式比较直接，用户可以直接通过网络问答社区邀请话题优秀回答者获得自身想要的答案，相对于在海量的网络资源中检索效率较高。

关注高质量用户是一种积极获取外界知识的途径。在“知乎”中用户关注高质量的用户后便可以在自身首页的动态中了解其他用户最新了解的知识，在一定程度上可以为用户带来许多新鲜的知识。通过关注来增强与外界接触也是“知乎”此类网络问答社区的主要功能之一，可以及时更新前沿知识。

回答高质量问题是用户外化自身隐性知识最好的途径之一。用户可以通过回答问题温故自身的旧知识，同时还可以锻炼自身的语言表达能力。此外，其他用户的问题也可以变相激发用户对于某一问题的认知，加深用户对问题的理解并赋予知识创新的机会，对知识自增长有着尤为重要的影响。

7 结 语

为了研究网络问答社区的知识传播机制和学习效用，本文选取“知乎”作为研究对象。获取“知乎”海量的用户信息，基于这些信息，遴选出 80 位意见领袖，并通过意见领袖之间的相互关注关系构建社会网络关系矩阵和社会网络关系图。结合 R 语言，分析此社会网络具备高聚集系数、低平均路径长度的小世界特征，并采用中间中心度、点度中心度、结构洞等指标衡量网络中的节点影响力。为了更好地分析节点的性质，本文使用相关性分析和层次聚类的方式降低数据维度，并通过熵权法加权求和的方式得到用户的知识能力指数和传播影响力指数；依据 Cowan 模型构建隐性知识传播动力学模型，并且结合用户的知识能力指数对用户知识传播和知识学习效果进行系统仿真。通过系统仿真结果发现：知识存量越少的节点在网络中的学习效率越高，而具备适当知识存量但存量依然较低

的节点在网络中的知识存量增长倍数最大, 本文将这种现象称之为知识的“海绵效应”。此外, 为了避免对知识的疲倦, 用户应及时更新关注网络中的新兴权力节点, 从而使得自身的学习效果处于更好的状态。为了保持良好的学习状态, 用户可以通过多提问题、多关注高质量用户和浏览他人的动态、多回答高质量问题来提升自身的知识量。

本文研究存在一些问题: 关于传播动力学模型的假设, 未考虑到网络动态的随机重连。未来可以通过对邻接矩阵数值的随机概率变化和增加新节点等方式将网络动态化。

参考文献:

- [1] 韦勇娇. 认知盈余时代, 网络问答社区的知识传播模式探析——以“知乎”为例[J]. 广西职业技术学院学报, 2014, 7(1): 30-34, 40. (Wei Yongjiao. Analysis of Modes of Knowledge Dissemination Online Q&A——A Case Study of Zhihu Q&A Website[J]. Journal of Guangxi Vocational and Technical College, 2014, 7(1): 30-34, 40.)
- [2] 吕峰. 基于隐马尔科夫模型的问答社区用户知识贡献意愿研究[D]. 哈尔滨: 哈尔滨工业大学, 2015. (Lv Feng. Research on the Knowledge Contribution Intention in Q&A Community Based on Hidden Markov Model[D]. Harbin: Harbin Institute of Technology, 2015.)
- [3] 金家华. 社会化问答社区中用户知识行为的影响因素研究[D]. 哈尔滨: 哈尔滨工业大学, 2015. (Jin Jiahua. Research on the Influence Factors of User's Knowledge Activities in Online Social Q&A Community[D]. Harbin: Harbin Institute of Technology, 2015.)
- [4] 陈娟, 高杉, 邓胜利. 社会化问答用户特征识别与行为动机分析——以“知乎”为例[J]. 情报科学, 2017, 35(5): 69-74, 80. (Chen Juan, Gao Shan, Deng Shengli. Characteristics Recognition and Motivation Analysis of Social Q&A Users——A Case Study of 'Zhihu' Website[J]. Information Science, 2017, 35(5): 69-74, 80.)
- [5] 黄梦婷, 张鹏翼. 社会化问答社区的协作方式与效果研究: 以知乎为例[J]. 图书情报工作, 2015, 59(12): 85-92. (Huang Mengting, Zhang Pengyi. Collaboration Forms and Effectiveness in Social Q&A Communities: A Case Study of Zhihu.com[J]. Library and Information Service, 2015, 59(12): 85-92.)
- [6] 刘佩, 林如鹏. 网络问答社区“知乎”的知识分享与传播行为研究[J]. 图书情报知识, 2015(6): 109-119. (Liu Pei, Lin Rupeng. Research on Knowledge Sharing and Dissemination Behavior of Online Q&A Services: Taking Zhihu as an Example[J]. Documentation, Information & Knowledge, 2015(6): 109-119.)
- [7] Liu Z, Jansen B J. Questioner or Question: Predicting the Response Rate in Social Question and Answering on Sina Weibo[J]. Information Processing & Management, 2018, 54(2): 159-174.
- [8] Jin J, Li Y, Zhong X, et al. Why Users Contribute Knowledge to Online Communities: An Empirical Study of an Online Social Q&A Community[J]. Information & Management, 2015, 52(7): 840-849.
- [9] Lou J, Fang Y, Lim K H, et al. Contributing High Quantity and Quality Knowledge to Online Q&A Communities[J]. Journal of the Association for Information Science and Technology, 2013, 64(2): 356-371.
- [10] 左遥, 梁英, 毕晓迪, 等. 社会化问答网站知识传播网络推断方法[J]. 计算机学报, 2018, 41(1): 82-97. (Zuo Yao, Liang Ying, Bi Xiaodi, et al. An Inference Method of Knowledge Diffusion Network in Community Question Answering Sites[J]. Chinese Journal of Computers, 2018, 41(1): 82-97.)
- [11] 宁菁菁. 基于“弱关系理论”的知识问答社区知识传播研究——以知乎网为例[J]. 新闻知识, 2014(2): 98-99, 50. (Ning Jingjing. Research on Knowledge Communication in Knowledge Q&A Community Based on “Weak Relationship Theory”——A Case Study of Zhihu Network[J]. News Research, 2014(2): 98-99, 50.)
- [12] Cowan R, Jonard N. Network Structure and the Diffusion of Knowledge[J]. Journal of Economics Dynamics and Control, 2004, 28(8): 1557-1575
- [13] Matlab [K/OL]. [2018-04-13]. <https://www.mathworks.com/>.
- [14] 王嘉. 思想政治教育视域下的网络意见领袖研究[D]. 大连: 大连理工大学, 2013. (Wang Jia. Researches on Network Opinion Leader for Ideological and Political Education [D]. Dalian: Dalian University of Technology, 2013.)
- [15] 赵基明, 邱均平, 黄凯, 等. 一种新的科学计量指标——h 指数及其应用述评[J]. 中国科学基金, 2008(1): 23-32. (Zhao Jiming, Qiu Junping, Huang Kai, et al. A New Scientometric Indicator——Review on h Index and Its Application[J]. Bulletin of National Natural Science Foundation of China, 2008(1): 23-32.)
- [16] Rstudio [K/OL]. [2018-04-13]. <https://www.rstudio.com/>.
- [17] 汪小帆, 李翔, 陈关荣. 网络科学导论[M]. 北京: 高等教育出版社, 2012: 233. (Wang Xiaofan, Li Xiang, Chen Guanrong. Network Science: An Introduction[M]. Beijing: Higher Education Press, 2012: 233.)
- [18] Tan P-N, Steinbach M, Kumar V. 数据挖掘导论(完整版)[M]. 范明, 范宏建等译. 北京: 人民邮电出版社, 2011: 38-39. (Tan P-N, Steinbach M, Kumar V. Introduction to Data Mining[M]. Translated by Fan Ming, Fan Hongjian, et al. Beijing: Post & Telecom Press, 2011: 38-39.)
- [19] 张继国, 辛格. 信息熵——理论与应用[M]. 北京: 中国水利水电出版社, 2012: 1-2. (Zhang Jiguo, Singh V P.

- Information Entropy—Theory and Application[M]. Beijing: China Water & Power Press, 2012: 1-2.)
- [20] 芦白茹, 张坤严, 心娥, 等. 基于 AHP 法和熵权法组合权重对 RC 结构危险房屋多层次模糊综合评定研究[C]. 见: 第 25 届全国结构工程学术会议论文集, 2016: 89. (Lu Bairu, Zhang Kuyan, Xin E, et al. Multi-hierarchy Fuzzy Comprehensive Assessment of RC Structure Dangerous Buildings Based on Combined Weights of AHP and Entropy-weighted Method[C]// Proceedings of the 25th National Conference on Structural Engineering. 2016: 89.)
- [21] Shannon C E. A Mathematical Theory of Communication[J]. Bell System Technical Journal, 1948, 27(3): 379-423.
- [22] Kermack W O, McKendrick A G. Contributions to the Mathematical Theory of Epidemics. II.—The Problem of Endemicity[J]. Proceedings of the Royal Society of London A, 1932, 138(834): 55-83.
- [23] 李中师. 基于个体差异的复杂网络下隐性知识传播模型研究[D]. 杭州: 浙江工业大学, 2012. (Li Zhongshi. Study of Tacit Knowledge Transfer Based on Individual Differences in Complex Social Networks[D]. Hangzhou: Zhejiang University of Technology, 2012.)
- [24] 张红兵. 虚拟企业知识转移的研究[M]. 北京: 经济管理出版社, 2009: 56. (Zhang Hongbing. Research on the Knowledge Transfer in Virtual Enterprise[M]. Beijing: Economics & Management Publishing House, 2009: 56.)
- [25] 汪应洛. 系统工程[M]. 第 4 版. 北京: 机械工业出版社, 2008: 82. (Wang Yingluo. System Engineering[M]. The 4th Edition. Beijing: China Machine Press, 2008: 82.)

作者贡献声明:

王忠义: 提出研究思路;
张鹤铭: 设计研究方案, 分析数据, 进行实验, 起草论文;
黄京: 采集、清洗数据;
李春雅: 论文最终版本修订。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据由作者自存储, E-mail: wzywzy13579@163.com。
[1] 王忠义, 张鹤铭. user-info.xlsx. 知乎 42 万用户原始数据。
[2] 王忠义, 张鹤铭. 80user.xlsx. 80 位意见领袖关系矩阵、特征描述和影响力分析数据。

收稿日期: 2018-03-16
收修改稿日期: 2018-05-29

Studying Knowledge Dissemination of Online Q&A Community with Social Network Analysis

Wang Zhongyi¹ Zhang Heming¹ Huang Jing² Li Chunya³

¹(School of Information Management, Central China Normal University, Wuhan 430079, China)

²(Wuhan Polytechnic, Wuhan 430074, China)

³(School of Business, Nantong Institute of Technology, Nantong 226002, China)

Abstract: [Objective] This paper analyzes the social network structure and knowledge dissemination mechanism of an online Q&A community, aiming to reveal the role of network nodes, and improve the learning efficiency. [Methods] First, we used the social network analysis and the entropy weight methods to describe the opinion leader's knowledge and influence. Then, we built a knowledge dissemination model based on the Cowan model for the Q&A community. Finally, we examined the internal knowledge learning results of the network through system simulation. [Results] I. The nodes with less knowledge had higher learning efficiency in the target network; II. The knowledge volumes of some nodes increased rapidly, while those of the nodes with larger knowledge stock increased slowly; III. The knowledge dissemination rate of this network has been decreasing; IV. There is strong correlation between knowledge increase and the index of knowledge and communication abilities. [Limitations] The dynamic random reconnection of network was not examined in this paper. [Conclusions] This paper offers practical advice to improve users' learning experience in the online Q&A community.

Keywords: Social Network Analysis Information Entropy Knowledge Dissemination Cowan Model Spongy Effect