



社会网络分析

案例：构建知网论文作者合作网

邵鹏 博士

副教授 硕士生导师

西安工程大学管理学院

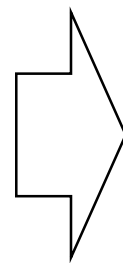
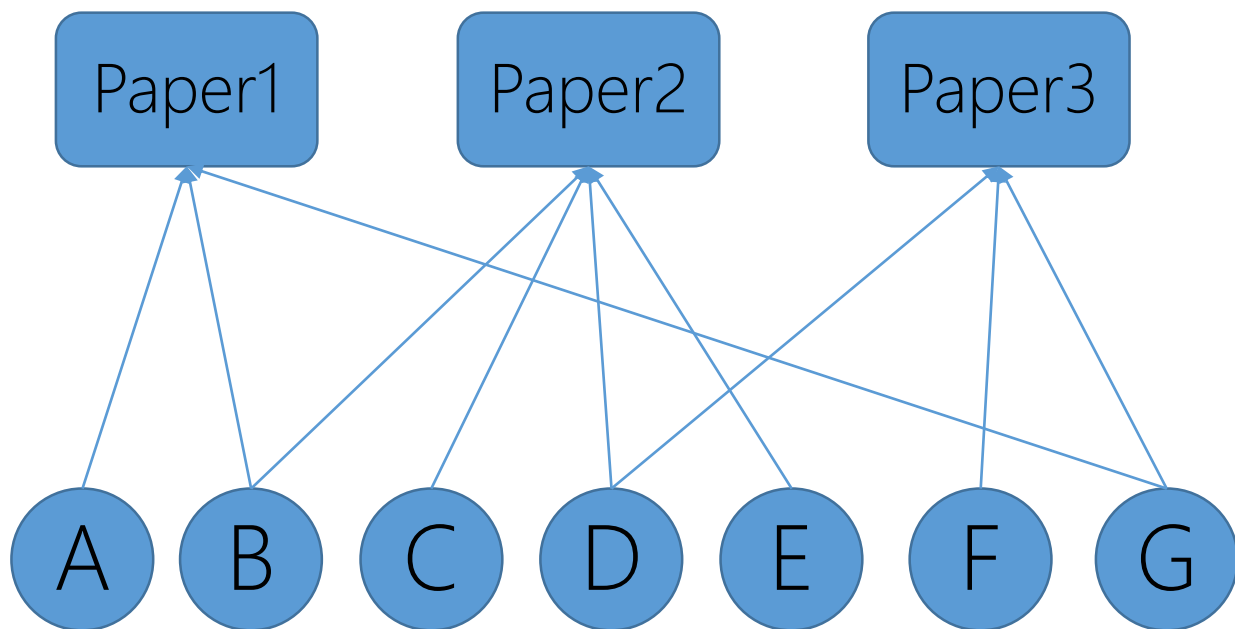
shaopengde@sohu.com

inebm.cn



问题

同一篇论文有好几个作者共同完成，某个作者完成过好多篇论文但每次合作的人不一样。如何构建论文合作网络？



	A	B	C	D	E	F	G
A		1					
B	1		1	1	1		
C		1		1	1		
D		1	1		1	1	1
E		1	1	1			
F				1			1
G				1		1	

1 导出文献基本信息（包括作者）



2 使用EXCEL/VBA清洗数据



3 使用VBA构建网络



4 使用Gehpi网络可视化



1 导出文献基本信息

以“西安工程大学管理学院为”单位
检索核心期刊文献377条

The screenshot shows the CNKI search interface. At the top, there are navigation tabs: 高级检索 (Advanced Search), 专业检索 (Professional Search), 作者发文检索 (Author Publication Search), 句子检索 (Sentence Search), 一框式检索 (One-box Search), 知识元检索 (Knowledge Element Search), and 引文检索 (Citation Search). The main search area includes a search box with the text "西安工程大学管理学院" (Xi'an Polytechnic University) and a dropdown menu for "作者单位" (Author Unit). Below the search box, there are several checkboxes for search options: 包含资讯 (Include Information), 网络首发 (Network First Issue), 增强出版 (Enhanced Publishing), 基金文献 (Fund Literature), 中英文扩展 (Chinese and English Extension), and 同义词扩展 (Synonym Extension). The "中英文扩展" checkbox is checked. There are also fields for "时间范围" (Time Range) and "来源类别" (Source Category). The "来源类别" section has several checked options: 全部期刊 (All Journals), SCI来源期刊 (SCI Source Journals), EI来源期刊 (EI Source Journals), 北大核心 (Beida Core), CSSCI, and CSCD. At the bottom of the search area, there are buttons for "重置条件" (Reset Conditions), "检索" (Search), and "结果中检索" (Search in Results). On the right side, there is a sidebar with the title "高级检索使用方法" (Advanced Search Usage) and a list of instructions for using advanced search operators. At the bottom of the page, there is a navigation bar with tabs for "总库" (Total Library), "中文" (Chinese), "外文" (Foreign), "学术期刊" (Academic Journals), "学位论文" (Theses), "会议" (Conferences), "报纸" (Newspapers), "年鉴" (Yearbooks), "图书" (Books), "专利" (Patents), "标准" (Standards), and "成果" (Achievements). The bottom right corner shows the search results: "共找到 377 条结果" (Found 377 results) and "1/8" (Page 1 of 8).



1 导出文献基本信息

导出至EXCEL

导出格式

• GB/T 7714-2015 格式引文

• 知网研学 (原E-Study)

• CAJ-CD 格式引文

• MLA 格式引文

• APA 格式引文

• 查新 (引文格式)

• 查新 (自定义引文格式)

• Refworks

• EndNote

• NoteExpress

• NoteFirst

GB/T 7714-2015 格式引文

预览

批量下载

导出

复制到剪贴板

打印

xls

doc

[1]张颖颖,舒朝阳,王进富,李浩.创业平台资源警觉、意义给赋与在解企业机会开发:多重中介模型检验[J/OL].科技进步与对策:1-8[2021-04-12].[http://dx.doi.org/10.1224.G3.20210506.1645.010.html](http://dx.doi.org/10.1224/G3.20210506.1645.010.html).

[2]王进富,李嘉辉,张颖颖,李浩.政府资助波动、技术市场发展程度与国有企业R&D投入[J/OL].科技进步与对策:1-8[2021-05-17].[http://dx.doi.org/10.1224.G3.20210506.1310.006.html](http://dx.doi.org/10.1224/G3.20210506.1310.006.html).

[3]和征,曲姣姣,李勃.基于MAS面向语义Web服务的绿色创新生态系统运作模型[J].现代电子技术,2021,44(09):92-98.

[4]吕英,王正斌,姚海博.女性董事、团体动力与企业社会责任——性别协同还是团体协同?[J].财经论丛,2021(04):94-105.

[5]郗飞,王铁山.物流业与制造业专业集聚、协同集聚与区域经济增长[J].企业经济,2021,40(04):88-97.

[6]郗飞.物流业与制造业协同集聚的空间关联与影响因素[J].统计与决策,2021,37(07):113-117.

[7]邵景峰,李宁,蔡再生.基于信噪比与综合赋权的涤纶低弹丝工艺优化模型[J].东华大学学报(自然科学版),2021,47(01):109-116.

[8]刘岩,沈聪,裴云龙,高艳慧.技术知识基础对企业成为关键研发者的影响研究:一个被中介的调节模型[J].科技管理研究,2021,41(05):61-67.

[9]王小红,范佳佳,程振.会计稳健性对非效率投资影响的实证研究——基于财务柔性的中介作用[J].财会通讯,2021(03):27-31.

[10]张岭,张田莉,张胜.风险投资对技术创新绩效的影响研究——基于政府干预的调节作用[J].科技管理研究,2021,41(03):100-108.

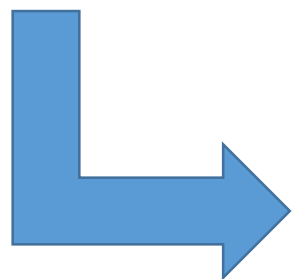


1 导出文献基本信息

Author-作者	Title-题名	Source-刊名	Year-年	Roll-卷	Period-期	PageCount-页码
张颖颖,舒朝阳,王进富,李浩	创业平台资源警觉、意义给赋与在孵企业机会开发:多重中介模型检验	科技进步与对策				1-8
王进富,李嘉辉,张颖颖,李浩	政府资助波动、技术市场发展程度与国有企业R&D投入	科技进步与对策				1-8
和征,曲姣姣,李勃	基于MAS面向语义Web服务的绿色创新生态系统运作模型	现代电子技术	2021	44	09	92-98
吕英,王正斌,姚海博	女性董事、团体动力与企业社会责任——性别协同还是团体协同?	财经论丛	2021		04	94-105
鄢飞,王铁山	物流业与制造业专业集聚、协同集聚与区域经济增长	企业经济	2021	40	04	88-97
鄢飞	物流业与制造业协同集聚的空间关联与影响因素	统计与决策	2021	37	07	113-117
邵景峰,李宁,蔡再生	基于信噪比与综合赋权的涤纶低弹丝工艺优化模型	东华大学学报(自然科学版)	2021	47	01	109-116
刘岩,沈聪,裴云龙,高艳慧	技术知识基础对企业成为关键研发者的影响研究:一个被中介的调节模型	科技管理研究	2021	41	05	61-73
王小红,范佳佳,程振	会计稳健性对非效率投资影响的实证研究——基于财务柔性的中介作用	财会通讯	2021		03	27-31
张岭,张田莉,张胜	风险投资对技术创新绩效的影响研究——基于政府干预的调节作用	科技管理研究	2021	41	03	100-108
李勃,王丛迪,和征	绿色产品协同创新中契约形式对供应商创新性的影响——供应商绿色创新资源优先配置的中介及产品复杂性与技术新颖性的调节	科技进步与对策				1-10
邵景峰,李宁,蔡再生	基于模糊多准则的涤纶低弹丝生产工艺参数优化	纺织学报	2021	42	01	46-52+58
贾丽娜,孙辉,曾涛,林晓文	证券投资基金共同持股、信息偏好与股价信息含量	金融理论与实践	2021		01	67-75



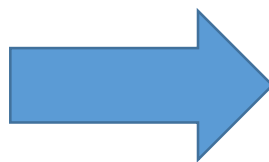
2 使用EXCEL/VBA清洗数据



关键词
分列

Author-作者			
张颖颖	舒朝阳	王进富	李浩
王进富	李嘉辉	张颖颖	李浩
和征	曲姣姣	李勃	
吕英	王正斌	姚海博	
鄢飞	王铁山		
鄢飞			
邵景峰	李宁	蔡再生	
刘岩	沈聪	裴云龙	高艳慧
王小红	范佳佳	程振	
张岭	张田莉	张胜	
李勃	王丛迪	和征	
邵景峰	李宁	蔡再生	
贾丽娜	孙辉	曾涛	林晓文
王宁	杨芮	周密	于玲玲
邵景峰	李宁	袁玉楼	
王渊	左温慧	田梦	

标记重名



粘贴到1列

1	张颖颖
2	王进富
3	和征
4	吕英
5	鄢飞
6	鄢飞
7	邵景峰
8	刘岩
9	王小红
10	张岭
11	李勃
12	邵景峰
13	贾丽娜
14	王宁
15	邵景峰
16	王渊
17	邵鹏
18	邵景峰
19	和征
20	刘岩
21	王小红
22	张岭
23	马创涛
24	陈恒

空行?
重复行?

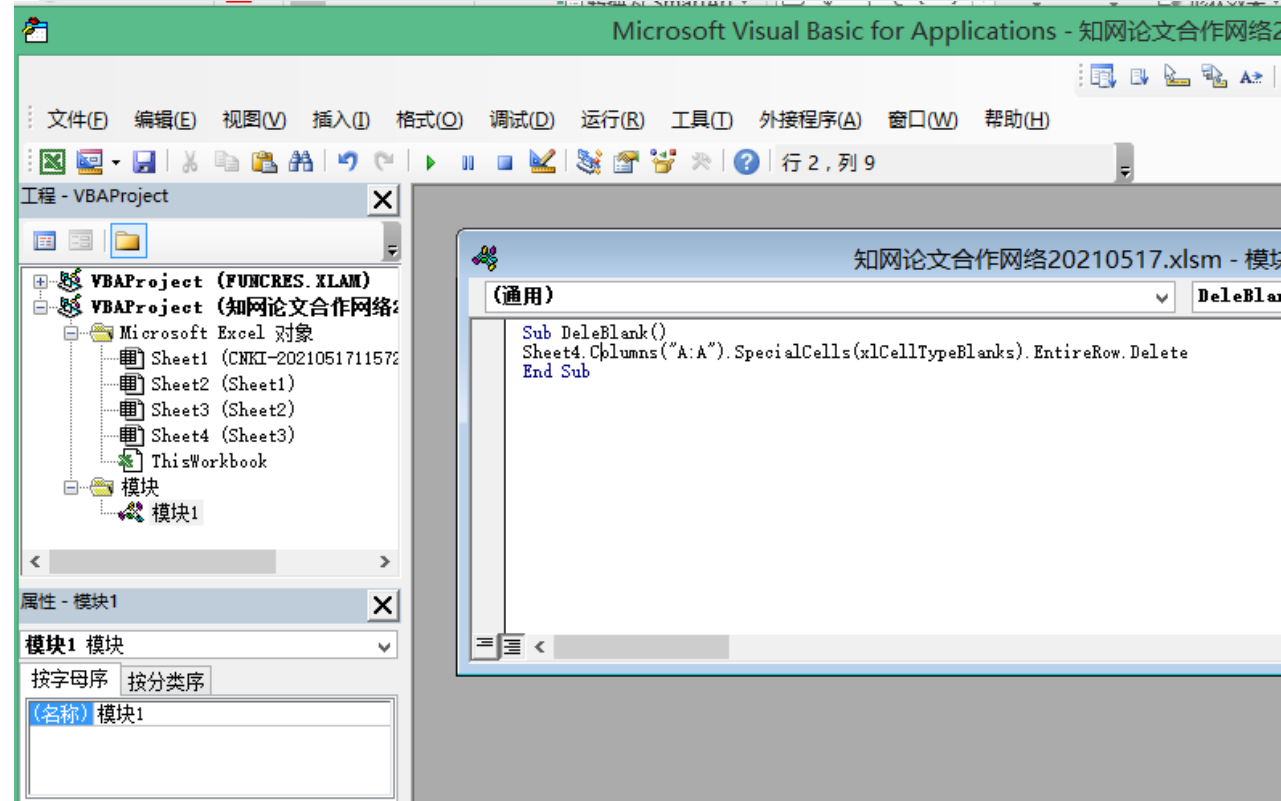
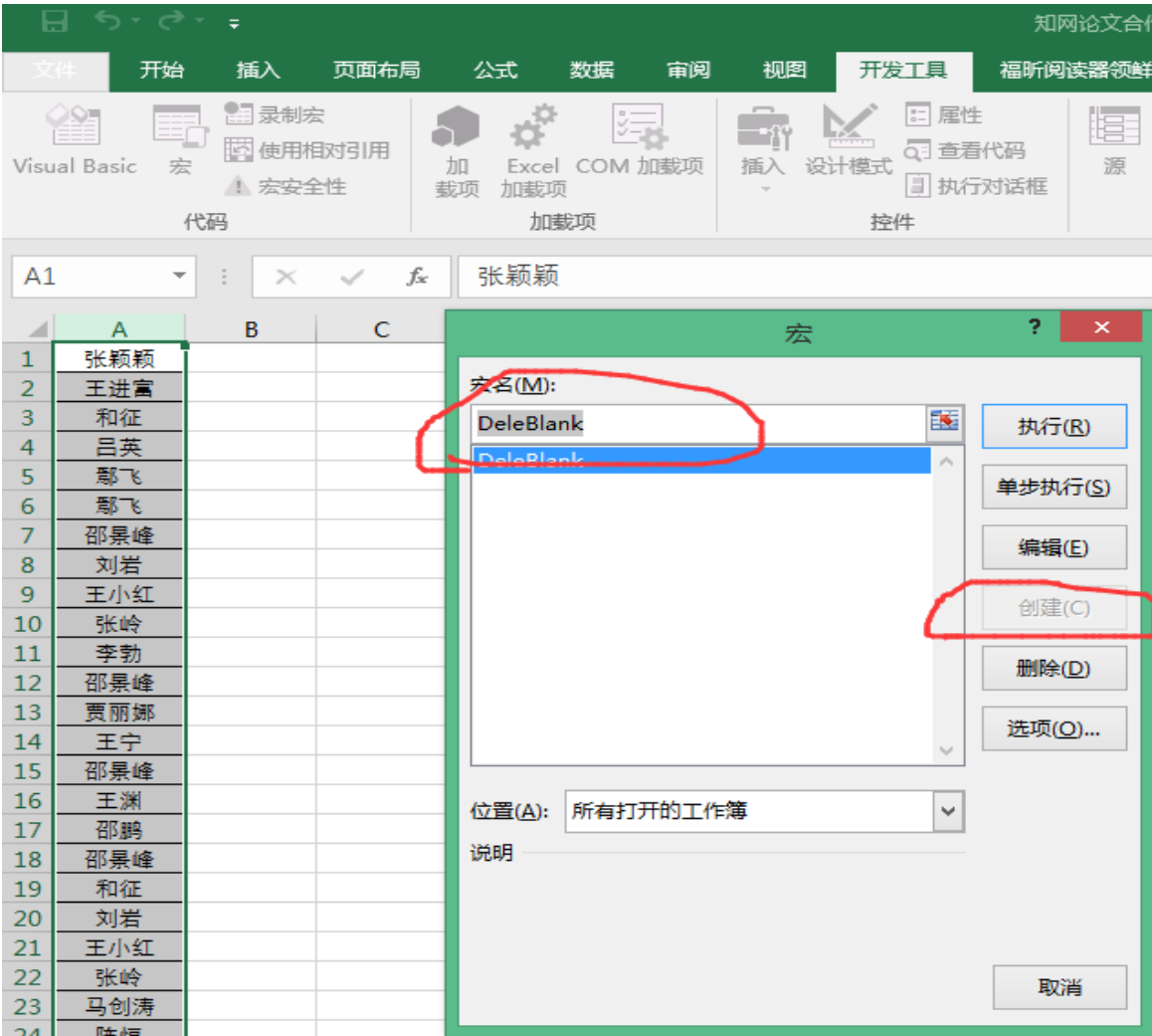


2 使用EXCEL/VBA清洗数据

EXCEL另存为xlsm格式，创建宏

使用VBA将第一列的空行删掉

```
Sub DeleBlank()  
Sheet4.Columns("A:A").SpecialCells(xlCellTypeBlanks).EntireRow.Delete  
End Sub
```





2 使用EXCEL/VBA清洗数据

使用VBA将第一列的重复的行删掉

```
Sub DeleRedun()  
Dim xRow As Integer  
Dim i As Integer  
xRow = Sheet4.Range("A65536").End(xlUp).Row  
For i = 1 To xRow  
    For j = i + 1 To xRow  
        If Sheet4.Cells(j, 1) = Sheet4.Cells(i, 1) Then  
            Sheet4.Rows(j).Delete  
            j = j - 1  
            xRow = xRow - 1  
        End If  
    Next  
Next  
End Sub
```



每一行的
作者名是
唯一的

	A	
31	陈宝东	
32	车明好	
33	张群林	
34	马贵梅	
35	吴芳	
36	胡忆楠	
37	孟亚丽	
38	张庆庆	
39	李艳2	
40	龚冬生	
41	庞芳兰	
42	胡海青	
43	白晓波	
44	阎亮	
45	张克英	
46	郑健	
47	王莹	
48	郝向举	
49	张延涛	
50	吴穹	
51	王保忠	
52	诸裕祥	
53	张正林	
54	彭再云	



3 使用VBA构建网络

新建一个sheet，准备构建矩阵

	A	B	C	D	E	F	G	H	I	J
1										
2	长颖颖									
3	王进富									
4	和征									
5	吕英									
6	鄂飞									
7	邵景峰									
8	刘岩									
9	王小红									
10	张岭									
11	李勃									
12	贾丽娜									
13	王宁									
14	王渊									
15	邵鹏									
16	马创涛									
17	陈恒									
18	梁杰									
19	贾艳玲									
20	单英骥									
21	李阳									
22	雷卫									
23	孙佳徽									
24	吕霄									

如果某两个作者同时是
同一篇论文的作者



则在两个作者的对应的
空格+1



3 使用VBA构建网络

对该左sheet，循环行，对于每一行的前4个作者

G	H	I	J
Author-作者			
张颖颖	舒朝阳	王进富	李浩
王进富	李嘉辉	张颖颖	李浩
和征	曲姣姣	李勃	
吕英	王正斌	姚海博	
鄢飞	王铁山		
鄢飞			
邵景峰	李宁	蔡再生	
刘岩	沈聪	裴云龙	高艳慧
王小红	范佳佳	程振	
张岭	张田莉	张胜	
李勃	王丛迪	和征	
邵景峰	李宁	蔡再生	
贾丽娜	孙辉	曾涛	林晓文
王宁	杨芮	周密	于玲玲
邵景峰	李宁	袁玉楼	
王渊	左温慧	田梦	



寻找在右sheet矩阵中的位置，并在对应位置+1

	A	B	C	D	E	F	G
1							
2	张颖颖		2				
3	王进富	2					
4	和征						
5	吕英						
6	鄢飞						
7	邵景峰						
8	刘岩						
9	王小红						
10	张岭						
11	李勃						
12	贾丽娜						
13	王宁						
14	王渊						
15	邵鹏						
16	马创涛						
17	陈恒						
18	梁杰						
19	贾艳玲						
20	单英骥						
21	李阳						
22	雷卫						



3 使用VBA构建网络

```
Sub GenMatrix()  
S2Row = Sheet2.Range("A65536").End(xlUp).Row  
S5Row = Sheet5.Range("A65536").End(xlUp).Row  
Dim string1, string2 As String  
Dim myrange1, myrange2 As Range  
For i = 1 To S2Row  
  For j1 = 7 To 9  
    For j2 = j1 + 1 To 10  
      string1 = Sheet2.Cells(i, j1)  
      string2 = Sheet2.Cells(i, j2)  
      If string1 <> "" And string2 <> "" Then  
        Set myrange1 = Sheet5.Columns("A:A").Find(what:=string1, lookAt:=xlWhole)  
        Set myrange2 = Sheet5.Columns("A:A").Find(what:=string2, lookAt:=xlWhole)  
        Sheet5.Cells(myrange1.Row, myrange2.Row) = Sheet5.Cells(myrange1.Row, myrange2.Row) + 1  
        Sheet5.Cells(myrange2.Row, myrange1.Row) = Sheet5.Cells(myrange2.Row, myrange1.Row) + 1  
      End If  
    Next  
  Next  
Next  
For S5 = 2 To S5Row  
  Sheet5.Cells(1, S5) = Sheet5.Cells(S5, 1)  
Next  
End Sub
```

循环行，对于每一行的前4个作者（7-10列）

寻找矩阵sheet中的位置，并在对应位置+1

将列作者名复制到行



3 使用VBA构建网络

另存为xlsx格式文件，
为导入Gephi准备

知网论文合作网络20210517.xlsm - Excel

数据 审阅 视图 开发工具 福昕阅读器领鲜版 Power Pivot 百度网盘 告诉我您想要做什么... 登录 共享

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1		张颖颖	王进富	和征	吕英	鄢飞	邵景峰	刘岩	王小红	张岭	李勃	贾丽娜	王宁	王渊	邵鹏	马创涛	陈恒	梁杰
2	张颖颖		7								1							
3	王进富	7		2			11				2			1				
4	和征		2				2				13							
5	吕英																	
6	鄢飞																	
7	邵景峰		11	2												4		
8	刘岩																	
9	王小红																	
10	张岭																	
11	李勃	1	2	13														
12	贾丽娜																	
13	王宁																	
14	王渊		1															
15	邵鹏																	3
16	马创涛						4											
17	陈恒																	
18	梁杰															3		
19	贾艳玲												1					
20	单英璞														3		1	
21	李阳																	
22	雷卫																	
23	孙佳徽																	
24	吕霄																	
25	王铁山					1												
26	郭文钰																	
27	刘晓喆																	
28	李丹																	
29	牛一凡																	

KNKI-20210517115725675 Sheet1 Sheet2 Sheet3 Sheet4



4 使用Gehpi网络可视化

Gephi 0.9.2

文件(F) 工作区 视图(V) 工具(T) 窗口(W) 帮助(H)

概览 数据资料 预览

数据表格 ×

节点 边

Spreadsheet (Excel)...

步骤

- Excel 常规选项
- 输入设置

Excel 常规选项 (1/2)

选择一个CSV文件输入:

F:\XPU\2021\2021SNA\2021SNA\案例:构建知网论文作者合作网\知网论文合作网络20210517.xlsx

表单: 导入数据

Sheet4 矩阵

预览:

	张颖颖	王进富	和征	吕英	鄢飞	邵景峰	刘岩	王小
张颖颖		7						
王进富	7		2			11		
和征		2				2		
吕英								
鄢飞								
邵景峰		11	2					
刘岩								

< 上一步(B) 下一步 > 完成(F) 取消 帮助(H)

节点: 374

边: 625

无向图

过滤

统计 ×

设置

平均加权重 9.872 运行 ?

网络直径 14 运行 ?

图密度 0.009 运行 ?

点击次数 运行 ●

模块化 0.85 运行 ?

PageRank 运行 ●

连接部件 运行 ●

节点概述

平均聚类系数 0.8 运行 ?

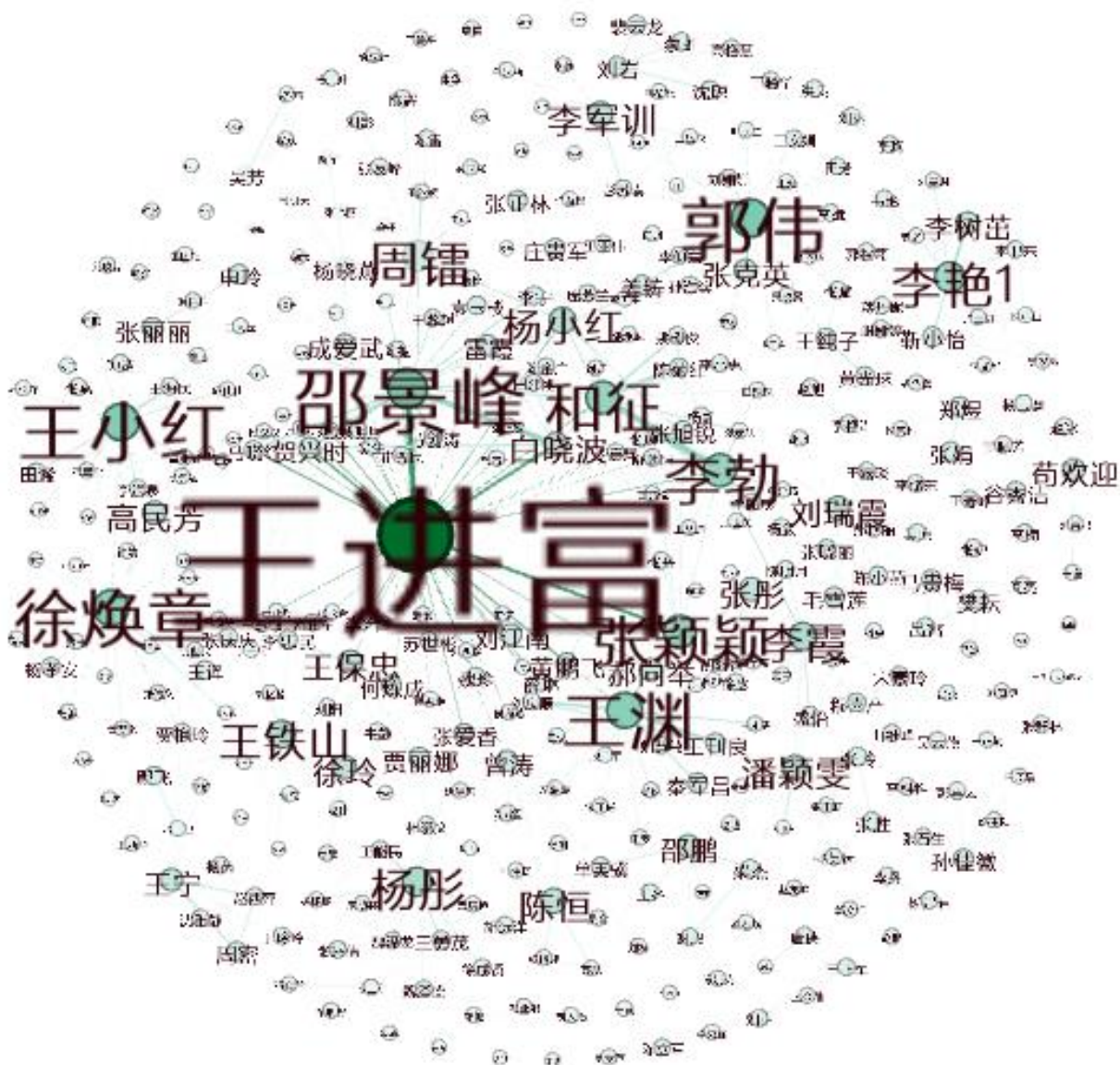
特征向量中心度 运行 ?

边概述

平均路径长度 5.761 运行 ?

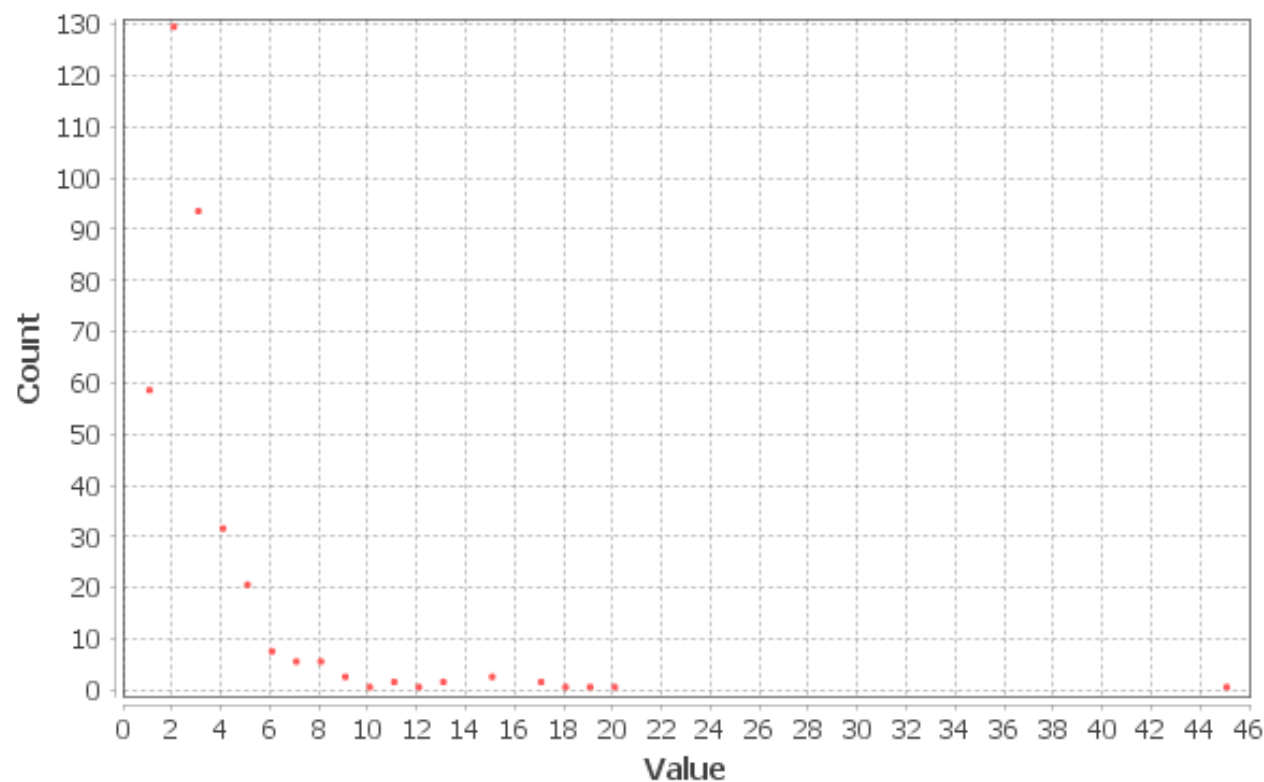


4 使用Gehpi网络可视化



Average Degree: 3.342

Degree Distribution





谢谢