



社会网络分析

案例：基于专利构建上市公司网络

邵鹏 博士

副教授 硕士生导师

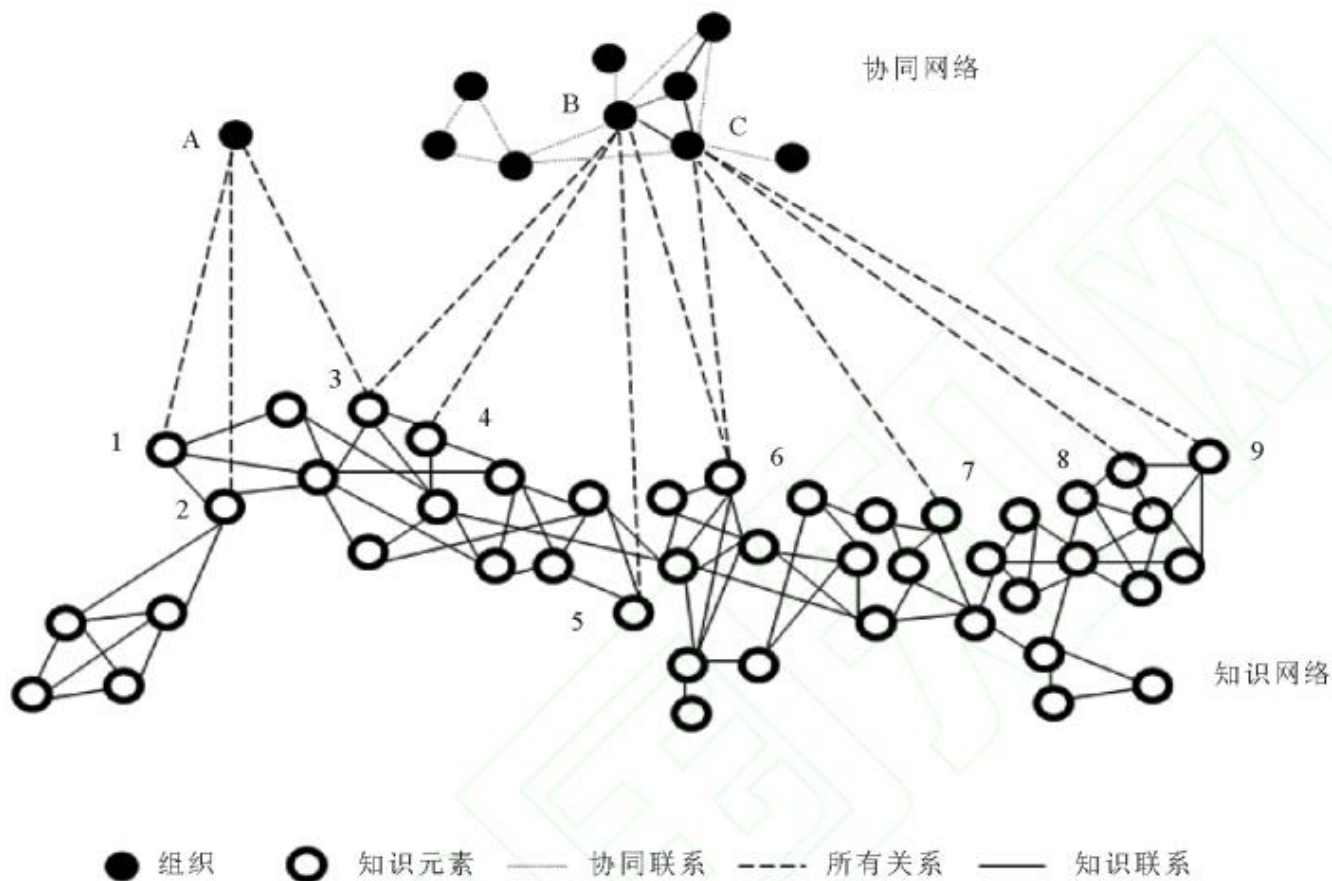
西安工程大学管理学院

shaopengde@sohu.com

inebm.cn

问题

专利可能由多个机构申请，多个专利可能面向同一个细分领域。如何基于专利的合作关系和行业隶属关系构建上市公司网络？



上方网络代表企业协同网络，每个网络节点代表一家企业，而纽带是指两个企业间的合作关系。

下方网络则代表知识网络，每个节点代表一个知识元素，纽带则代表知识元素间通过创新成果进行联合应用。

1 获取上市公司数据



2 按照时间窗整理数据



3 基于专利构建公司网络



4 其他网络



1 获取上市公司数据

确定行业

选择2012证监会行业分类中的“计算机、通信和其他电子设备制造业”。



代码选择

代码/简称搜索 可选代码 [0]

- 专用设备制造业
- 汽车制造业
- 铁路、船舶、航空航天和其它运输设备制造业
- 电气机械及器材制造业
- 计算机、通信和其他电子设备制造业
- 仪器仪表制造业
- 其他制造业
- 废弃资源综合利用业
- 金属制品、机械和设备修理业



筛选条件2011年前上市的、非ST,*ST企业。导出上市公司代码EXCEL，一共有N家企业？



1 获取上市公司数据

获取财务数据

将代码导入国泰安，获取这N家公司2011-2020年10年的相关指标数据。

	A	B	C	D	E	J
1	code					
2	000002					
3	000003					
4	000011					
5	000012					
6	000013					
7	000015					
8	000016					
9	000017					
10	000018					
11	000019					
12	000020					
13						
14						
15						

CSMAR | 上市公司研究服务平台

搜索指标

- 基本信息
 - + 公司文件
 - + 特色指标
 - + 治理综合信息
 - + 三会基本信息
- 股本股东
 - + 股本信息
 - + 股东信息
- 行情指标
- 业绩预告
- 财务报表
 - + 资产负债表
 - + 利润表
 - + 现金流量表
- 财务附注
 - + 应收帐款
 - + 预付账款
 - + 其他应收款
 - + 存货
 - + 开发支出
 - + 研发投入
 - + 无形资产
 - + 商誉
 - + 固定资产、累计折旧、减值准备
 - + 借款抵押
 - + 应付职工薪酬
 - + 主营业务收入
 - + 营业收入
 - + 其他业务收入
 - + 政府补助
- 财务分析



1 获取上市公司数据

获取专利数据

用企业名称（包括子公司、分公司名称）在专利数据库（如INNOJOY）检索其在2007-2020年的发明专利。N个公司，则导出N个EXCEL数据。

重新检索-高级

申请日

选中本页

1 意图识别方法、装置
中国发明专利 ★★

申请号:CN20201164506
公开(公告)号:CN1127651
申请(专利权)人:[广东];
圳);深圳哈工大科技创新产
发明(设计)人:张民;肖康;卢
innojoy同族数:1 权项数:1
本发明公开一种意图识别
识别方法包括步骤:获取
向量输入双向LSTM模型
量计算文字信息中所有词
实现了不需更换数据转换

检索历史: < > ≡

AN/申请号	PNM/公开号	TI,ABST,CLM/名称,摘要,权利要求书	TI,ABST/名称,摘要
TI/名称	ABST/摘要	CLM/权利要求书	DESCR/说明书
SIC/分类号	PA/申请人	PATMS/申请人集合	INN/发明人
ADDR/地址	CAS/当前专利权人	AD/申请日	PD/公开日
CLS/当前法律的状态	CLSD/当前法律状态公告日	AGC/专利代理机构	NEWDLJG/代理机构集合
DAN/分案原申请号	LV/有效性	SQGGR/授权公告日	IAN/国际申请
			IPN/国际公布

下载 保存 监控

and or not () xor adj equ/10 xor/10 pre/10 % ? To 收起

2 一种运动教练姿势评估指导的实现方法和系统[ZH] 审中
中国发明专利 ★★

申请号:CN202011644232.2 申请日:2020.12.31
公开(公告)号:CN112784699A 公开(公告)日:2021.05.11
申请(专利权)人:[广东]; 康佳集团股份有限公司
发明(设计)人:袁义德;穆罕默德·阿德南;李涛;杨波;梁兴伟 优先权: 国际公布:
innojoy同族数:1 权项数:10
本发明公开一种运动教练姿势评估指导的实现方法和系统。用于采集用户姿势图像,
采用姿势模型提取所述用户姿势图像中的用户姿势关键点,将所述用户姿势关键点与
预先存储的教练姿势关键点进行比较,确定所述教练姿势关键点和所述用户姿势关键
点位置差与阈值,并确定用户姿势的标准度。本发明提供了一种可以随时随地接受专
业运动教练姿势评估指导的实现方法,一方面可以节省人们去健身房等路程时间,随

代理机构 深圳市君胜知识产权代理事务所(普通合伙) 4
268

一案双申

摘要
[中文] 本发明公开一种意图识别方法、装置、设备及计算机可读存储
介质,其中,所述意图识别方法包括步骤:获取文字信息,并通过语
嵌入表将文字信息转化成词向量;将词向量输入双向LSTM模型,通
双向LSTM模型的编辑层输出文本向量;根据文本向量计算文字信息
所有词汇的概率;将概率中的最大值对应的词汇确定为意图词汇,显
现了不需更换数据转换规则的情况下智能识别用户的意图,提高了识



1 获取上市公司数据

描述专利基本情况

哪些企业的专利最多？

名称	2020专利数	2019专利数	2018专利数
公司A			
公司B			
...			
公司N			

哪些IPC类下的专利最多？

名称	2020占比	2019占比	2018占比
IPC 1			
IPC 2			
...			
IPC m			



1 获取上市公司数据

描述两类专利数据

探索式创新，企业在第t年的探索式创新为企业第t年成功申请的技术专利且在前5年都未申请过。

利用式创新，企业在第t年申请的专利总量减去探索式创新绩效数值。

计算每个公司每年的**创新数量**、**探索式创新数量**、**利用式创新数量**。

公司A

	当年专利数	前5年各专利的代码前4位	当年专利的代码	探索式创新专利数	利用式创新专利数
2020					
2019					
2018					
2017					
...					
2011					



2 按照时间窗整理数据

专利研究通常并不一仅用当年的数据，而是通过时间窗的方式将5年或3年的数据进行整理。

以5年时间为窗口构建网络，如2011年用2007-2011的数据，2012年用2008-2012的数据。

以3年时间为窗口构建网络，如2011年用2009-2011的数据，2012年用2010-2012的数据。

本案例以5年时间窗为例整理数据。



2 按照时间窗整理数据

每个公司按时间窗整理

对于每个公司导出的专利EXCEL,建立2011-2020共10个sheet。

- sheet2011里面包括2007-2011的数据
- sheet2012里面包括2008-2012的数据
- sheet2013里面包括2009-2013的数据
- sheet2014里面包括2010-2014的数据
- sheet2015里面包括2011-2015的数据
- sheet2016里面包括2012-2016的数据
- sheet2017里面包括2013-2017的数据
- sheet2018里面包括2014-2018的数据
- sheet2019里面包括2015-2019的数据
- sheet2020里面包括2016-2020的数据

对于这10个sheet，第1-2列插入该公司的上市公司代码和公司名称。其中，上市公司代码用于“知识网络构建”中“公司-专利”隶属关系向“公司-公司”网络的转换。




2 按照时间窗整理数据

所有公司数据按时间窗汇总

建立2011-2020共10个EXCEL文件，每个EXCEL包括所有公司该区间的专利数据。

- EXCEL2011中，将所有公司sheet2011的数据粘在一起
- EXCEL2012中，将所有公司sheet2012的数据粘在一起
- EXCEL2013中，将所有公司sheet2013的数据粘在一起
- EXCEL2014中，将所有公司sheet2014的数据粘在一起
- EXCEL2015中，将所有公司sheet2015的数据粘在一起
- EXCEL2016中，将所有公司sheet2016的数据粘在一起
- EXCEL2017中，将所有公司sheet2017的数据粘在一起
- EXCEL2018中，将所有公司sheet2018的数据粘在一起
- EXCEL2019中，将所有公司sheet2019的数据粘在一起
- EXCEL2020中，将所有公司sheet2020的数据粘在一起



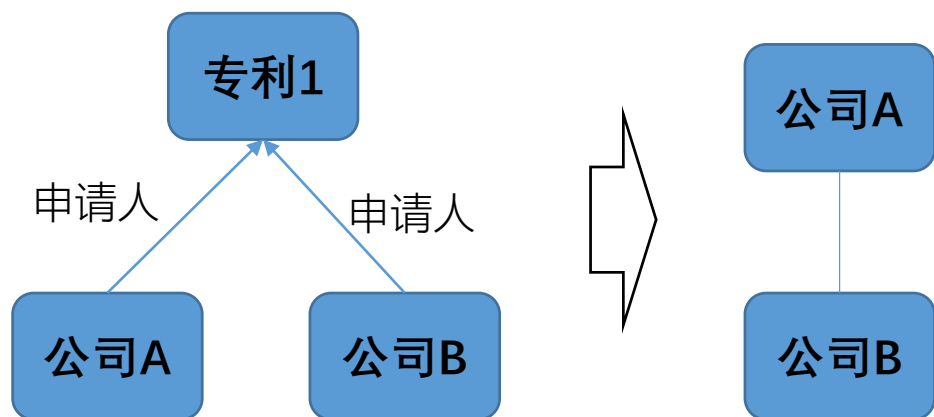
将这10个EXCEL另存为10个xlsm文件，为VBA编程构建网络准备。

3 基于专利构建公司网络

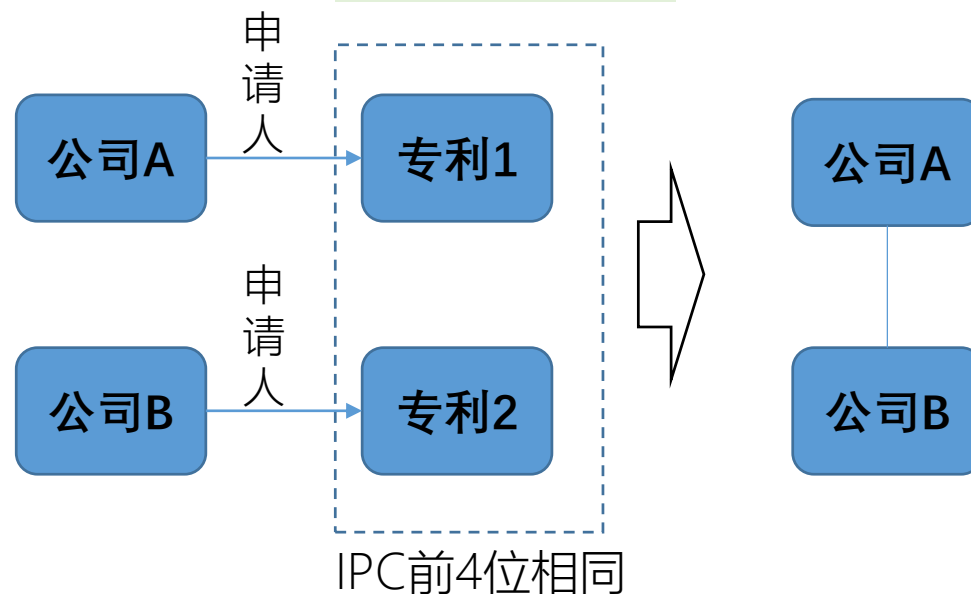
网络构建思路

参考“双层网络嵌入下结构洞对企业二元创新的影响研究”、“协同网络中知识域耦合对企业二元创新的影响”，由企业间合作关系形成协同网络，由知识元素组成知识网络。

协同网络



知识网络





3 基于专利构建公司网络

协同网络构建

将“申请人”列按照“；”进行分列，得到专利合作邻接数据。

U	V	W	X
申请人			
康佳集团 股份有限公司	苏州大学	哈尔滨工 业大学(深 圳)	深圳哈工 大科技创 新产业发 展有限公 司
康佳集团 股份有限公司			
康佳集团 股份有限公司	深圳市格 灵人工智 能与机器 人研究院 有限公司	哈尔滨工 业大学(深 圳)	深圳哈工 大科技创 新产业发 展有限公 司

具体代码参照“案例：构建知网论文作者合作网”，如果两个机构出现在同一行，则建立两个结构的邻接关系。

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	李勤	
吕英	王正斌	姚海博	
郇飞	王铁山		
郇飞			
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	
李勤	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	郇飞						
7	邵景峰						
8	刘岩						
9	王小红						
10	张岭						
11	李勤						
12	贾丽娜						
13	王宁						
14	王渊						
15	邵鹏						
16	马剑涛						
17	陈恒						
18	梁杰						
19	贾艳玲						
20	单英骥						
21	李阳						



3 基于专利构建公司网络

知识网络构建

提取“IPC”的前4位，如果两个公司在该时间窗拥有相同类型的专利（IPC前4位相同），则建立两个公司的连接。

M	N
主IPC分类号	IPC前4位
G06F40/216(2020.01)	G06F
G06K9/00(2006.01)	G06K
G06K9/00(2006.01)	G06K

具体代码参照“案例：基于连锁董事构建上市公司网络”，如果“IPC”前4位相同则建立两公司的邻接关系。

如果董事ID相同，则建立这两公司的邻接关系



```
n = Sheet1.Range("A65536").End(xlUp).Row
t = 1
For i = 3 To n
  For j = 2 To i - 1
    If Sheet1.Cells(j, 4) = Sheet1.Cells(i, 4) Then
      Sheet2.Cells(t, 1) = Sheet1.Cells(j, 1)
      Sheet2.Cells(t, 2) = Sheet1.Cells(i, 1)
      t = t + 1
    End If
  Next j
Next i
```



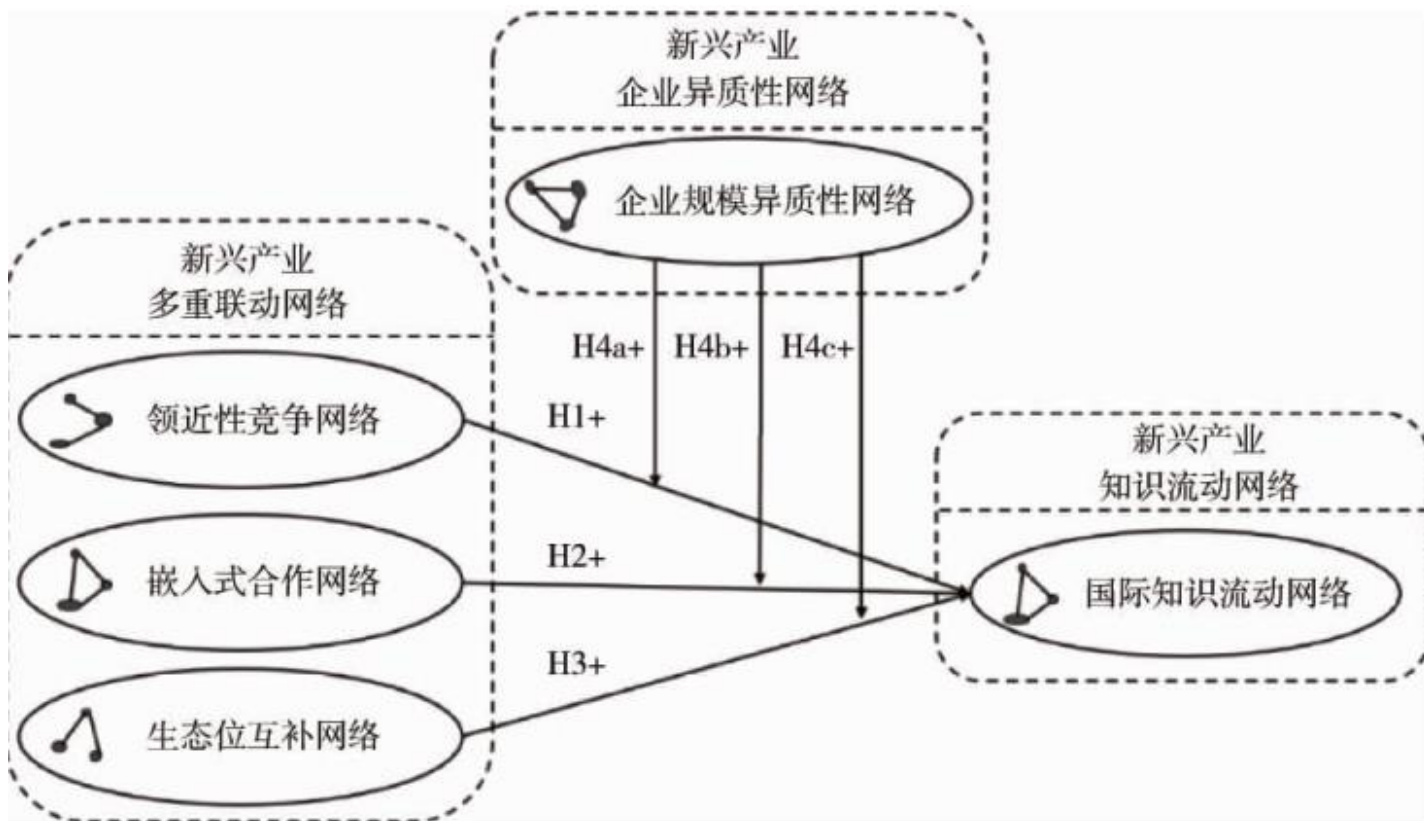
4 其他网络

许冠南,王丽明,周源.新兴产业多重联动网络对知识流动网络影响机制[J].科学学研究,2021,39(03):463-470.

伴随组织间关系网络化、生态化、国际化的发展趋势,新兴产业知识流动影响机制研究亦需从聚焦二元组织距离向关注组织间联动网络转变。为此,本文基于全球创新生态系统视域,提出新兴产业**竞争、合作、互补**的**多重联动网络**对知识流动网络影响机制概念模型,并对2005-2016年间全球3D打印产业开展实证研究。研究发现,新兴产业**邻近性竞争网络、嵌入式合作网络、生态位互补网络**均对**国际知识流动网络**有显著正向影响,**规模异质性网络**正向调节邻近性竞争网络与国际知识流动网络之间的关系。



4 其他网络



八个网络

公司信息

规模异质性网络
 研发强度异质性网络
 盈利能力异质性网络
 企业地理距离网络

邻近性竞争网络
 嵌入式合作网络
 生态位互补网络
 国际知识流动网络

专利信息



4 其他网络

表 1 变量测度及文献依据

	变量	测度方式	文献依据	数据来源
因变量	国际知识流动网络	专利引用矩阵	Jaffe 等, 1993; Jiang 等, 2019	DPCI
	邻近性竞争网络	技术领域相似性矩阵	Marrocu 等, 2013; Azagra – Caro & Consoli, 2016	DWPI
自变量	嵌入式合作网络	专利合作矩阵	郭建杰等, 2019	DWPI
	生态位互补网络	产业链互补矩阵	谢建国等, 2015	DWPI
	研发强度异质性网络	研发强度差距矩阵	Aristei 等, 2016; Sugheir 等, 2012	OSIRIS
控制变量	盈利能力异质性网络	利润率差距矩阵	原毅军、于长宏, 2019	OSIRIS
	企业规模异质性网络	员工规模差距矩阵	Lahiri, Nandini; Narayanan, Sriram 2013	OSIRIS
	企业地理距离网络	地理距离矩阵	Jaffe 等, 1993; Peri, 2005	谷歌地图



4 其他网络

国际知识流动网络

通过全球专利引用来表征，通过企业间的专利引用关系构建企业间专利引用矩阵，进而形成以企业为节点，以企业之间专利引用量为边的社会网络，以此表征国际知识流动网络[5][27]。其中专利引用矩阵中的元素为企业间专利引用量，取值越大，表示知识流动越频繁。

- [5] 许冠南,胡伟婕,周源,张舵.全球新兴制造网络知识流动影响要素及演化[J].科学学研究,2019,37(12):2194-2202.
- [27] Shann-Bin Chang, Kuei-Kuei Lai, Shu-Min Chang.Exploring technology diffusion and classification of business methods: Using the patent citation network,Technological Forecasting and Social Change,Volume 76, Issue 1,2009,Pages 107-117,



4 其他网络

邻近性竞争网络

借助企业间的技术领域相似性来表示其邻近性竞争关系，用企业间申请专利行业代码的余弦相似度来测度[28]，据此来构建邻近性竞争网络矩阵。该矩阵为多值矩阵，元素取值范围为 $[0,1]$ ，越接近1表示邻近性竞争越激烈。

[28]Emanuela Marrocu, Raffaele Paci, Stefano Usai.Proximity, networking and knowledge production in Europe: What lessons for innovation policy?,Technological Forecasting and Social Change,Volume 80, Issue 8,2013,Pages 1484-1498,



4 其他网络


余弦相似度文献

knowledge base may exchange information and knowledge and learn from each other more easily. To measure the technological, or cognitive, proximity across regions we compute a similarity index between region i and region j , based on the distribution of patenting activity among 44 sectors,⁹ defined as:

$$t_{ij} = 1 - \left(\frac{1}{2} \sum_{k=1}^{K=44} |l_{ik} - l_{jk}| \right)$$

where l_{ik} is the sectoral share of sector k in region i . The index t_{ij} is defined between zero (perfect dissimilarity of the sectoral distribution) and one (perfect similarity); thus, the higher the index value, the more similar the technological structure of the two regions and the higher the probability that they can exchange knowledge. The index is computed for each couple of

$$TP_{ij} = 1 - \left(\frac{1}{2} \sum_{k=1}^{k=NIPC} |t_{ik} - t_{jk}| \right)$$

- 
- t_{ik} 表示公司*i*在*k*分类的专利数占*i*公司专利总数之比， t_{jk} 表示公司*j*在*k*分类的专利数占I公司专利总数之比。
 - NIPC是IPC分类的个数，对*i*和*j*两家公司每个IPC分类的下的专利差异的绝对值进行求和。
 - 计算企业*i*和*j*申请专利行业代码的余弦相似度 TP_{ij} (邻近性竞争关系)。



4 其他网络

嵌入式合作网络

通过企业间合作申请专利来表征企业间嵌入式合作[23]，进而构建嵌入式合作网络矩阵。矩阵中元素为企业之间的专利合作申请频次，数值越大表示嵌入式合作越深入，该矩阵为对称矩阵。

[23]Guannan Xu, Yuchen Wu, Tim Minshall, Yuan Zhou. Exploring innovation ecosystems across science, technology, and business: A case of 3D printing in China, Technological Forecasting and Social Change, Volume 136, 2018, Pages 208-221,



4 其他网络

生态位互补网络

借助贸易互补指数的计算公式[29]，用企业所处产业价值链的位置互补来考量企业间的生态位互补关系。通过企业间所处价值链环节的余弦相似度来表征企业间的生态位同质化程度，用“1 - 生态位同质化程度”表示企业间的生态位互补关系，据此来构建生态位互补网络矩阵。矩阵元素的取值范围为 $[0,1]$ ，越接近 1 表示生态位互补程度越高。

[29]谢建国,杨海燕.互补还是替代——中美贸易竞争关系的测度与分析[J].国际贸易问题,2015(11):62-72.



4 其他网络

企业规模异质性网络

用企业员工数表征企业规模，进而用企业间员工数的自然对数之差来测度企业规模异质性，据此构建企业规模异质性网络矩阵。矩阵中元素取值越大，表示企业间规模异质性越大。

Size矩阵(企业规模异质性网络)单元格取值：

$$\text{Size}_{ij} = |\ln(i\text{公司员工数}) - \ln(j\text{公司员工数})|$$



4 其他网络

研发强度异质性网络

以企业研发支出与营业收入之比来计算研发强度，进而用企业研发强度自然对数之差来衡量企业间研发强度异质性，据此构建研发强度异质性矩阵。

RDS矩阵(研发强度异质性网络)单元格取值：

$$RDS_{ij} = |\ln(i\text{公司}RD) - \ln(j\text{公司}RD)|$$



4 其他网络

盈利能力异质性网络

以企业利润率为测度盈利能力，进而用企业利润率的自然对数之差来衡量企业间盈利能力异质性，据此构建盈利能力异质性矩阵。

Profit矩阵(盈利能力异质性网络)单元格取值：

$$\text{Profit}_{ij} = |\ln(i\text{公司利润率}) - \ln(j\text{公司利润率})|$$



4 其他网络

地理距离网络

以企业注册地之间的地理距离来计算企业地理距离[2]，据此构建企业地理距离矩阵。

Jaffe A B, Trajtenberg M, Henderson R. Geographic localization of knowledge spillovers as evidenced by patent citations [J]. The Quarterly Journal of Economics, 1993, 108(3) : 577-598 .



谢谢