

由中国人民大学信息资源管理学院冯惠玲教授、加小双副教授主讲的《数字人文导论》是中国人民大学本科生“数字人文荣誉辅修学位”系列课程之一，选课学生涵盖本、硕、博三类。课程内容包括数字人文导论（理论讲解、案例分享）和数字人文项目设计（实践操作）。

本系列推文共12期，是该课程学生课堂展示的成果报道。第七期由信息资源管理学院2021级博士生李少建和徐红霞共同完成，以伦敦玛丽女王大学Ruth Ahnert教授、剑桥大学Sebastian Ahnert教授以及波茨坦大学kim albrecht博士主持的“The Tudor Network”项目为研究对象，对该项目的背景、内容、成果、特色等进行系统梳理，最后对该项目进行评析，总结对今后数字人文项目的启示。

## The Tudor Network

# The Tudor Network

## The Tudor Network

### ◆ 项目简介 ◆

#### What

基于英国国家档案馆**123,850**封书信档案信息的数据化形成的跨越近**100**年，涉及**20,424**人的书信网络的可视化项目

#### Why

历史档案的数字化提供了从全新的角度审问历史资料的机会

目标是展示历史学术研究如何能够从大规模的网络分析、文本挖掘和相关的定量方法中受益

#### Who



Professor Ruth Ahnert (承担历史分析和数据清理工作)  
 BA MPhil PhD in Cambridge (剑桥本硕博)  
 Queen Mary University of London (任职伦敦玛丽女王大学)  
 Professor of Literary History & Digital Humanities (文学史&数字人文教授)  
 Interested in book history and epistolary culture (研究兴趣为书史与书信文化)



Professor Sebastian Ahnert (承担网络分析和数据整理工作)  
 University of Cambridge (任职剑桥大学)  
 Department of Chemical Engineering and Biotechnology (化学工程与生物技术系教授)  
 研究兴趣主要有两个，其一是从算法信息论的角度研究生物学中的结构和功能复杂性，以及算法信息论的跨学科应用；其二是网络分析，尤其是网络分析在生物学、人文学科和社会科学中的跨学科应用,还对基因组统计,布尔网络,自然语言处理,高斯过程和分子动力学感兴趣



Dr. Kim Albrecht (承担创意方向和数据可视化)  
 BA in graphic design and an MA in interface design (图形设计学士,界面设计硕士)  
 Ph.D. at the University of Potsdam in media theory (波茨坦大学媒体理论在读博士)  
 Interested in visual knowledge in post-digital age (研究兴趣为后数字时代视觉知识)

该项目得到英国艺术与人文研究委员会、斯坦福大学人文中心和伦敦玛丽女王大学资助

## How

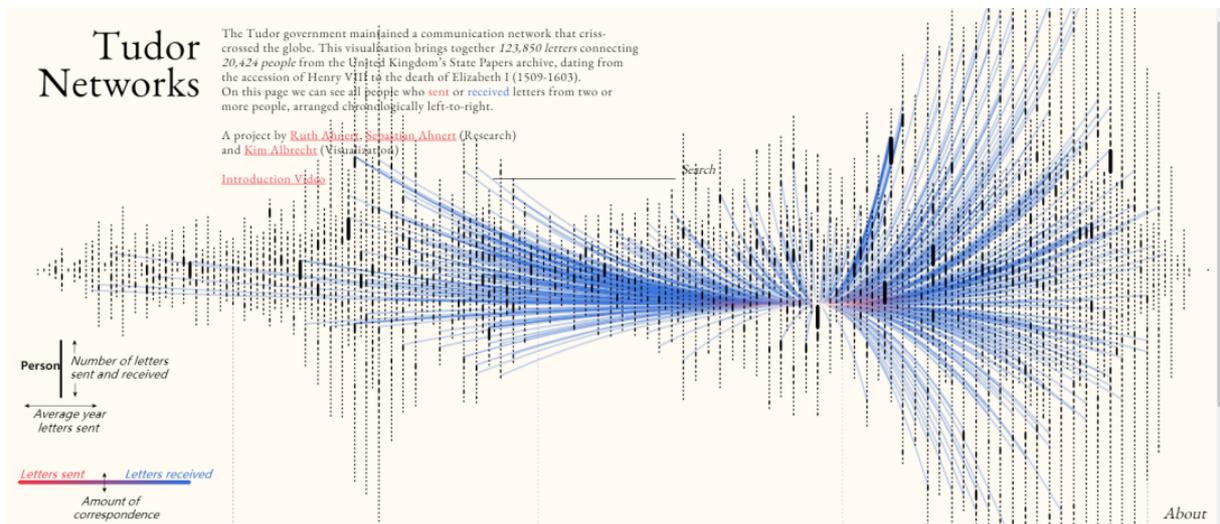


图1 The Tudor Network主界面随机截图

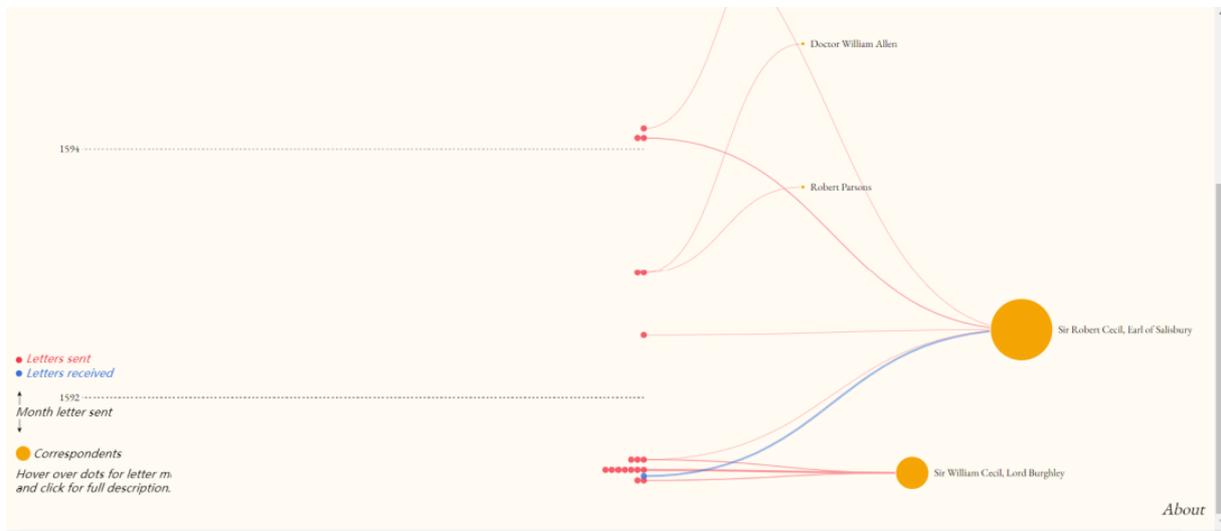


图2 按时间显示某人收发信图



图3 按地点显示某人收发信图

## ◆ 项目成果 ◆

### ▷ 网站

<http://tudornetworks.net/>

### ▷ 书籍

A book, Tudor Networks of Power.

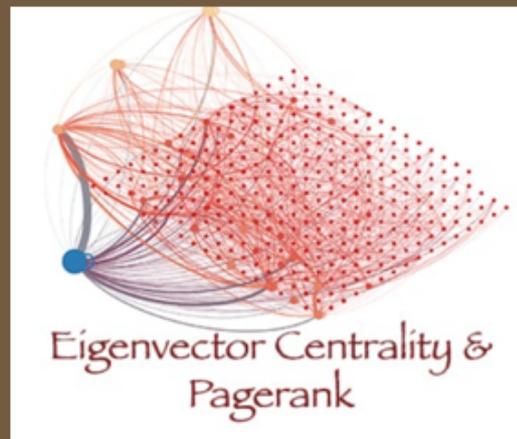
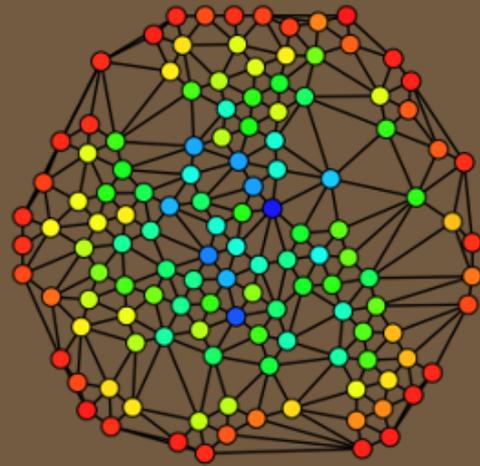
### ▷ 论文

Ruth Ahnert, Sebastian E Ahnert, Metadata, Surveillance and the Tudor State, History Workshop Journal, Volume 87, Spring 2019, Pages 27–51.

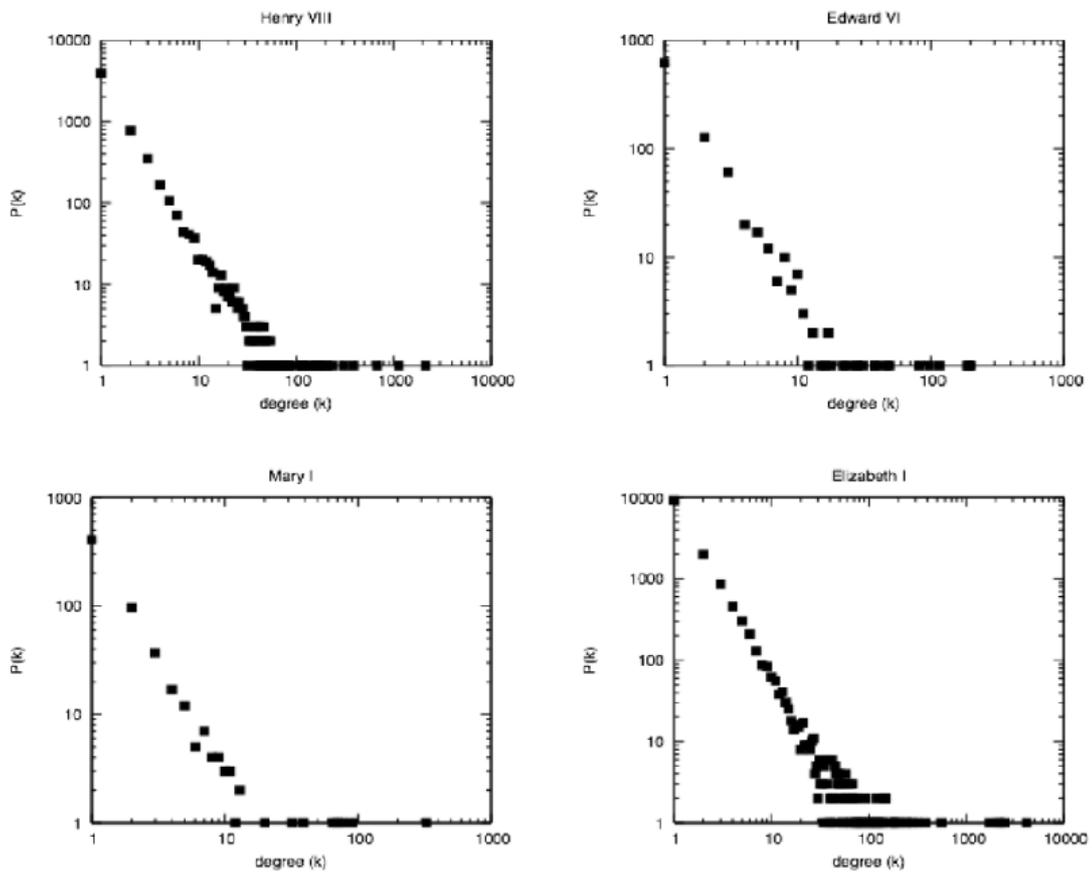
## Case Analysis

## 分析指标:

- Nodes (结点)
- Edges (边)
- Degree (度)
- In-degree (入度)
- Out-degree (出度)
- Betweenness (介数中心值)
- Strength (收发信总数)
- In-Strength (收信数)
- Out-strength (发信数)
- Eigenvector Centrality (本征向量中心)



## ▷ Degree



Figures 1a-d: These four graphs chart the distribution of the number of correspondence connections (known formally as 'degree distribution') for the reigns of Henry VIII, Edward VI, Mary I and Elizabeth I

图4 Degree分布图

▷排序分析：亨利八世时期，最高度的人是Thomas Cromwell (2149；首席秘书、枢机大臣、财政大臣)，之后是国王(1134)，Cardinal Thomas Wolsey(682；枢机主教)

▷长尾分布：在亨利八世时期，5785个结点中0.3%有100多条边。2.2%超过20条边，90.3%的边少于5条，1条边的结点占到68.1%。呈现长尾分布

▷幂律分布：数据点分布广泛，图b-d中中度为1的收发信方比例分别是67.8%、67.5%、66%

## ▷ Betweenness

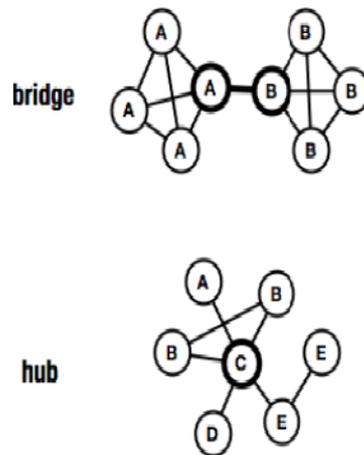


Figure 2: Two 'toy' networks illustrating why both hubs and bridges benefit from the measure of betweenness. In both cases, the shortest path between any nodes bearing different letters (e.g. A and B, or B and D) necessarily passes through the red nodes, because there are no other possible paths.

图5 Bridge&Hub作用示意图

- ▷桥梁: 例如外交官 (betweenness较高, degree较低)
- ▷中心/枢纽: 例如间谍 (betweenness较高, degree较低)

## ▷The Correlation between Degree and Betweenness

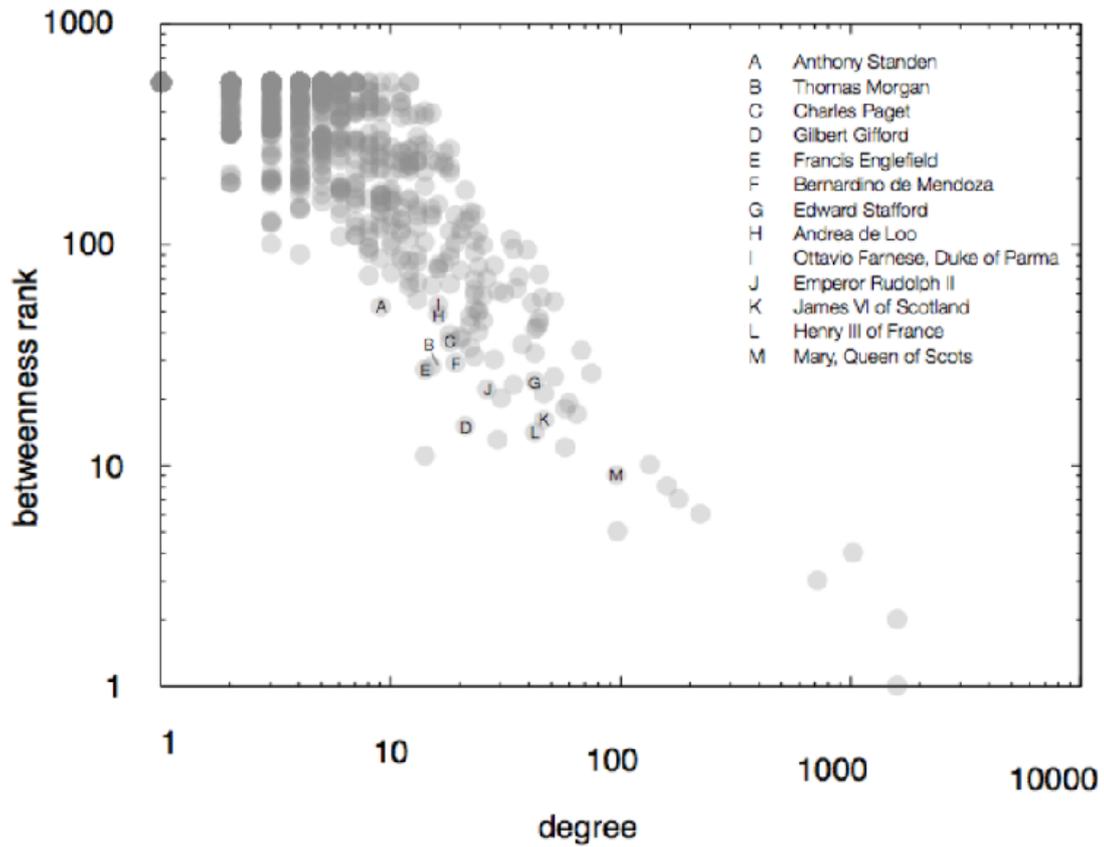


Figure 3: This graph charts degree against the betweenness ranking of each person who corresponded in the 1580s. The data points on the graph are varied in their opacity to signal the number of nodes with that score: the darker, the higher the concentration of people.

图6 16世纪80年代度与中介中心值的相关性分析图

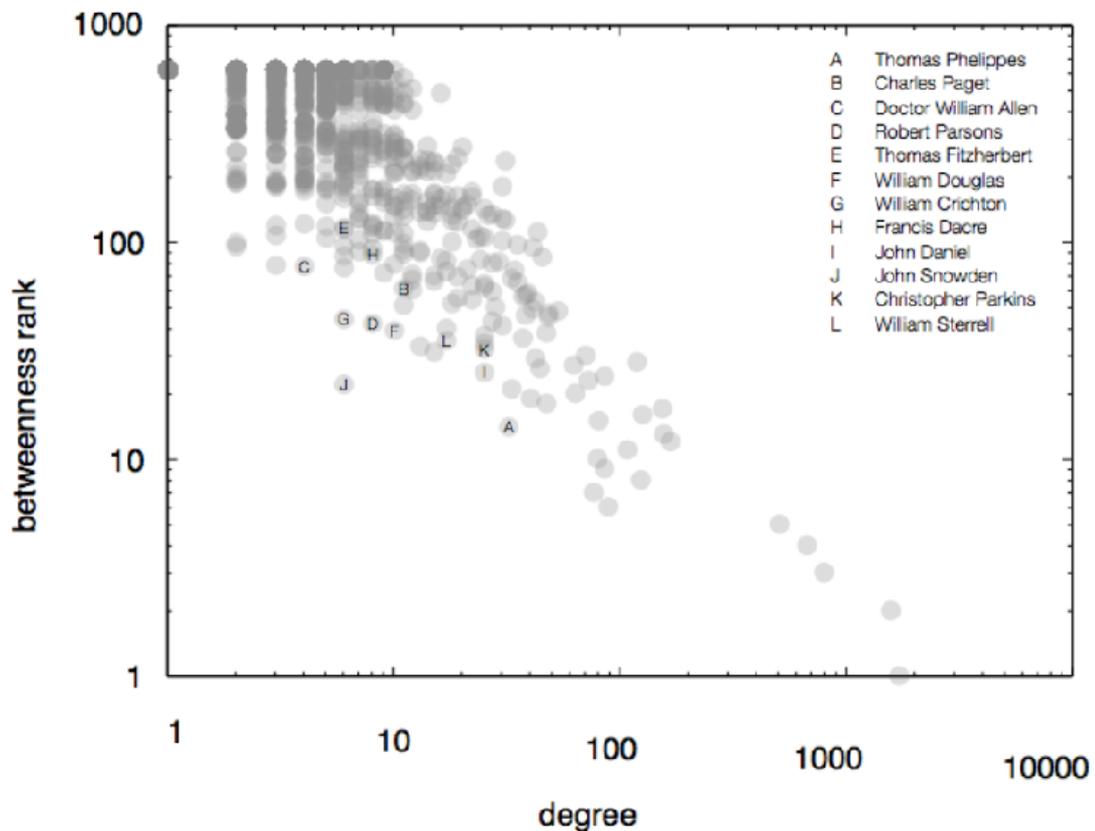


Figure 4: This graph charts degree against the betweenness ranking of each person who corresponded in the 1590s.

图7 16世纪90年代度与中介中心值的相关性分析图

▷左上角与右下角：图6共有5040人，左上角代表了3378（67%）的人，他们的betweenness为0，只与一个人通信。右下角度为1579，betweenness排序首位，是Francis Walsingham，他是首席秘书之一，以使用间谍网络闻名，所以成为了关键的信息中心

▷整体趋势线：不同时期趋势线相似，呈现结点度越高，介数中心值也越高的趋势（书信往来越频繁的人往往越重要）

▷异常值分析：位于趋势线上方和下方存在一些异常点。下方的异常点具有degree相对betweenness较低的特点。例如ABCDEF均有天主教信仰，以及均为阴谋者或双重间谍

### ▷定性与定量相结合的分析方法案例：双面间谍 John Snowdon

John Snowdon原名John Cecil，是一名英国天主教徒，毕业于牛津大学Trinity学院。1587-1588年的大部分时间里担任主教Allen的拉丁秘书。Allen和Parsons是伊丽莎白时期最主要的两个敌人。1591年，在Person的指挥下，Snowden和John Fixer作为间谍被派往英国，但他们的船被拦截，他们被囚禁审问。所以

Snowdon提供了Allen和Parsons计划的事情，并自荐为国王的秘密特工。为了赢得国王信任，Snowden在**信中写下了**西班牙队英格兰的主要做法，以及他自己的想法，提出了一个解决办法，既能成为忠实臣名，又能实践天主教信仰，成了**双面间谍**。所以他很重要，但是仍然只与少数关键领导交流，类似于**信息的守门人**，有多重路径为他提供了身份的保护

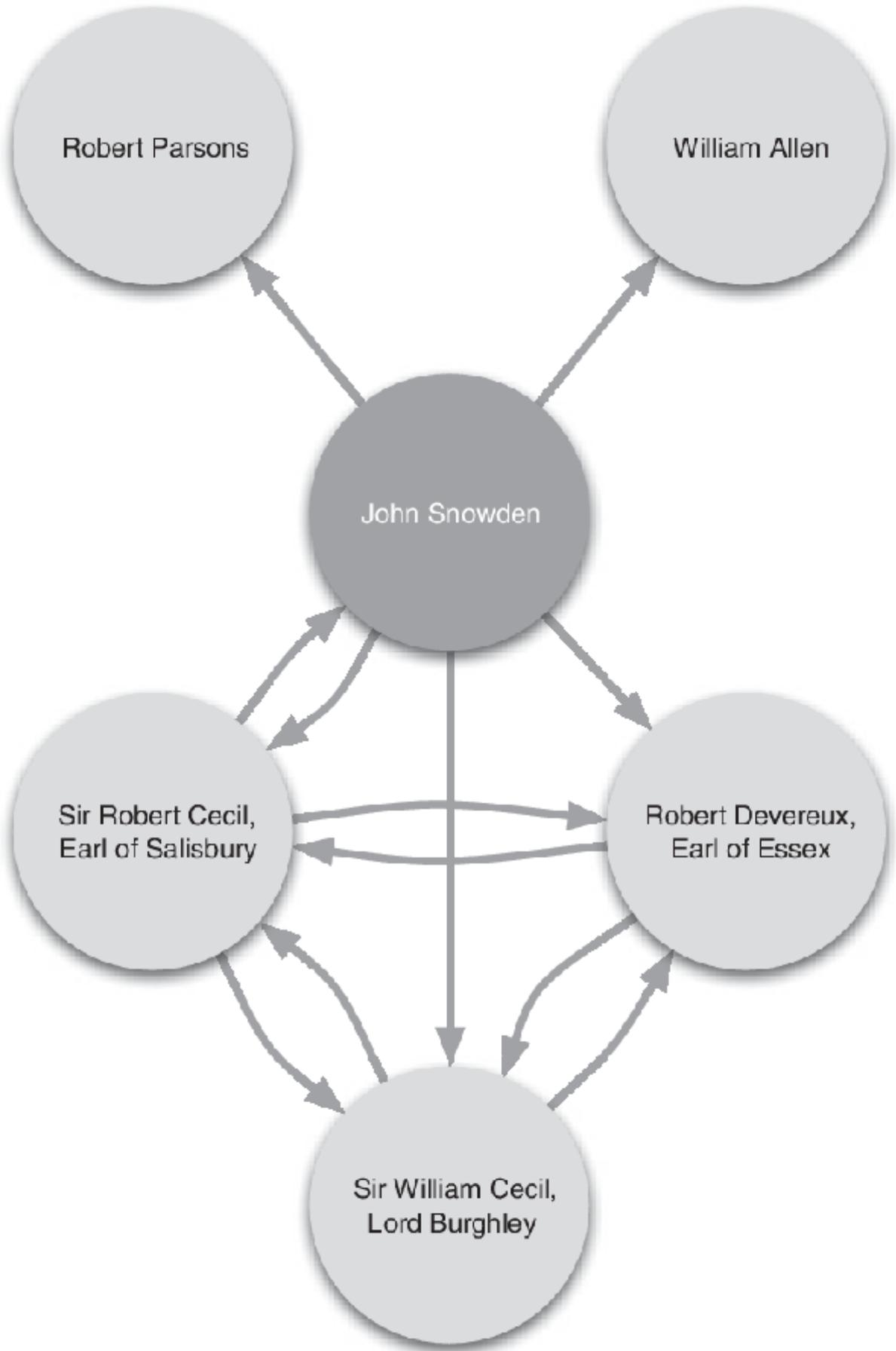


图8 John Snowden微观通信网络示意图

▷ **Eigenvector Centrality**

与君主或首席秘书有连接的将会有高的Eigenvector Centrality分数，为指示与**中央政府的距离**的粗略指标

## ▷ 总体相似度

### ▷ 相似度排序

例如与主教Allen最相似的15个人有：1) William Douglas, Earl of Angus, 2) Robert Persons, 3) Anthonio de Guarás, 4) Sir Gelli Meyrick, 5) Francis Dacre, 6) Anthony Babington, 7) Juan de Idiáquez, 8) Mathew Carnsew, 9) Pedro de Zubiaur, 10) Hugh Owen, 11) Thomas Paget, Lord Paget, 12) Sir William Catesby, 13) Gilbert Curll, 14) Charles of Lorraine, Duke of Mayenne, 15) Francesco de Mendoza

8人为来自不列颠群岛的天主教阴谋者，5人为天主教徒；其中4人是西班牙外交人员和军事领袖。他们都对英国的安全构成了威胁，信件大多数是由拦截而进入档案馆的。剩余的2人，4是同谋，8不属于这一类

### ▷ 发现特定活动的特定通信方式群体

1) Tommaso spinelli，是英国最早的常驻大使之一，曾在Margaret of Austria宫廷中任职。与他相似的20个人中除了1人以外，都在这一期间执行过外交任务

2) Pietro Bizzarri没有清晰的类别，他寄出很多信件，但是Eigenvector Centrality是0（他只接受一个人的信，这个人不给他人写信）。与他最相似的15个人中，有4个军事领导人和4个外交官。除此之外还有代理人、牧师和商人。除1人外，其余均参与了向政府的报告。与双面间谍不同的是其strength比较高。他们向女王的首席秘书发送了大量信件，却只有5位有记录

## 方法价值

通过网络指纹分析，理解群体及其随时间发生的变化

- 1) 计算个人网络指纹的相似性，发现群体及其特征
- 2) 发现政府监视所留下的指纹，发现某个特定时期内，政府主要关注的对象

## ◆ 项目特色 ◆

### ▷ 题材与数据来源

- (1) 题材独特
- (2) 数据来源可靠

### ▷ 呈现方式

- (1) 通信网络可视化呈现形式

## (2) 独特时空双维度呈现

### ▷资源重构

根据档案元数据建构通信网络

### ▷数据挖掘

为档案元数据的挖掘提供了一种新视角。基于大量书信档案得到分析结果，可利用人物之间的通信关系探寻到平时很难发觉的隐秘特征，帮助研究者充分利用书信档案中的信息

### ▷主创人员

学科背景多样

## ◆ 项目评析 ◆

### ▷项目可视化

- (1) 可视化粒度较粗，在空间维度上，无法获取准确的地理位置信息
- (2) 空间维度展示没有地图嵌入
- (3) 可视化界面较复杂

### ▷项目功能

- (1) 只能通过人名检索，检索方式单一
- (2) 没有数据下载API

### ▷项目维护

缺乏后期维护和更新，缺少未来迭代方向

### ▷用户体验

不流畅，易用性有待增强，缺少用户手册

### ▷项目建设

缺少公开的项目建设过程记录，不能让后来者站在前人的肩膀上

## ◆ 总结 ◆

### ▷提供高阶功能的数字人文项目

不仅可以提供查找、浏览等基础功能，更能够通过复杂网络等方法提供数据挖掘等高阶功能

▷元数据的分析价值

大量元数据构成的复杂网络可能会产生新的价值

▷数字人文项目中资源建构

对资源进行多维度统计和可视化，为繁杂的数据呈现了整体网络结构形式；对资源进行时空维度的展现能够发现新的价值

▷支持定量和定性相结合研究的数字人文项目

有时空维度的统计，支持定量的分析。有书信的原文，支持定性的分析

编辑：李少建 徐红霞

排版：李少建 徐红霞



公众号账号：rucdh2019

网址：<http://dh.ruc.edu.cn>

邮箱：[rucdh@ruc.edu.cn](mailto:rucdh@ruc.edu.cn)



## 中心简介

中国人民大学数字人文研究中心集人民大学多学科优势，秉持融合文理、协同创新之理念，开展数字人文理论研究、实践探索、人才培养和学术交流。