

Principles of Infographics

–Seminar–

Yasemin Aslan)

SPRU (Science Policy Research Unit)
Business School
University of Sussex



Week 7: 11 March 2022

Learning Outcomes

Learning outcome	Assessment mode
1 Explain the concept of network and list the main network indicators	ESS
2 Describe and apply the major techniques for the collection of network data and their statistical analysis	ESS, GPN + GWS
3 Identify the main characteristics of networks by means of network measures	ESS, GPN + GWS
4 Employ network analysis techniques to produce network data-based infographics	GPN + GWS

Note: ESS: Essay; GPN: Group Presentation; GWS: Group Written Submission

Overview

- 1 Network visualisation [recap]
- 2 Introduction to Gephi

Network visualisation [recap]

Network visualisation [recap]

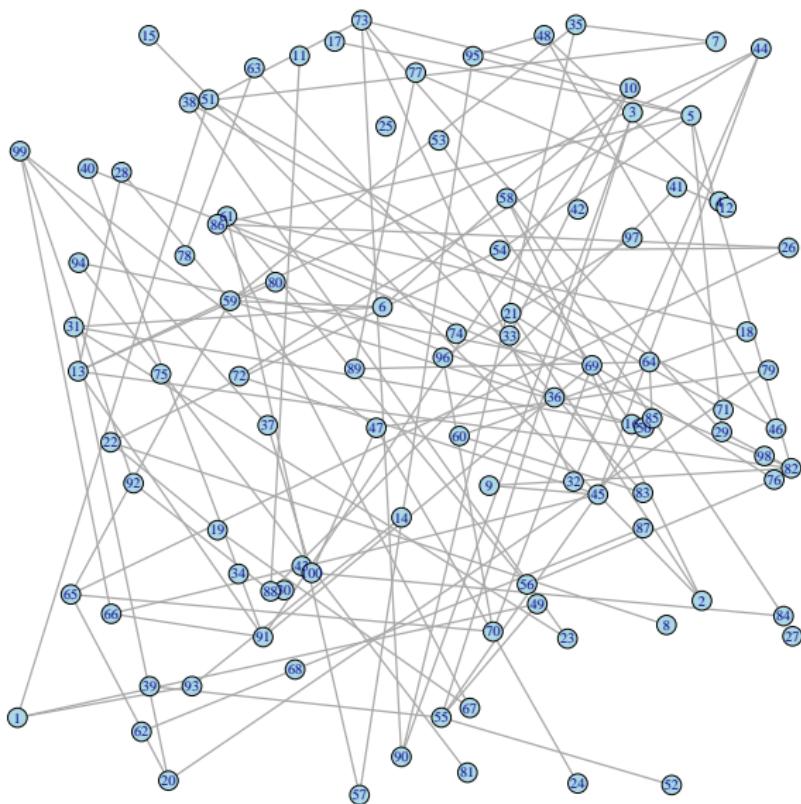
Layout

Let's generate a network $G(N, E)$ where:

- $N = 100$
- $x_{ij} = 1$ with $p = 0.02$

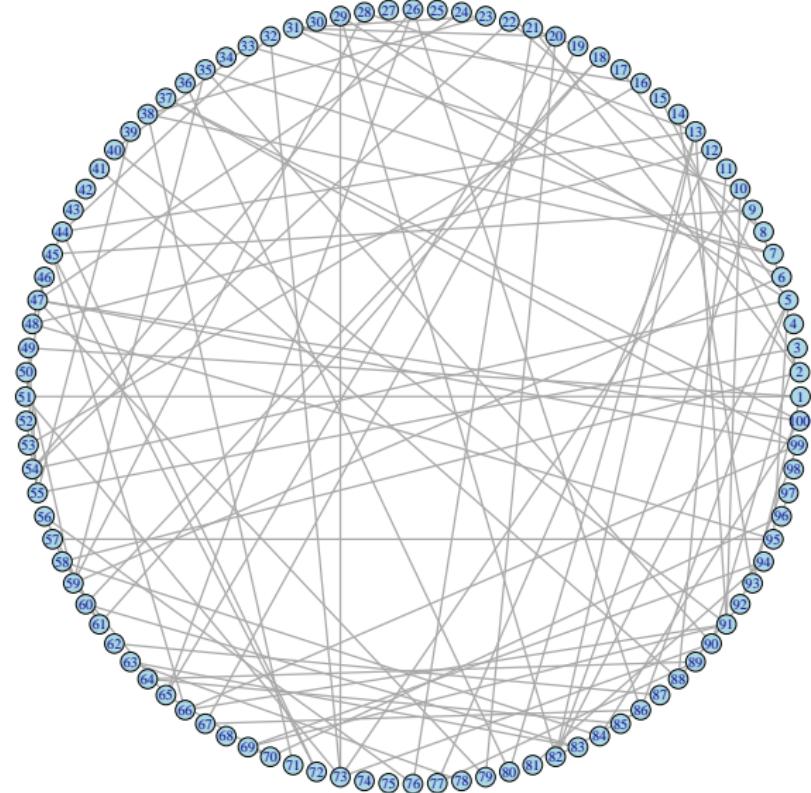
Network visualisation [recap]

Layout: Random



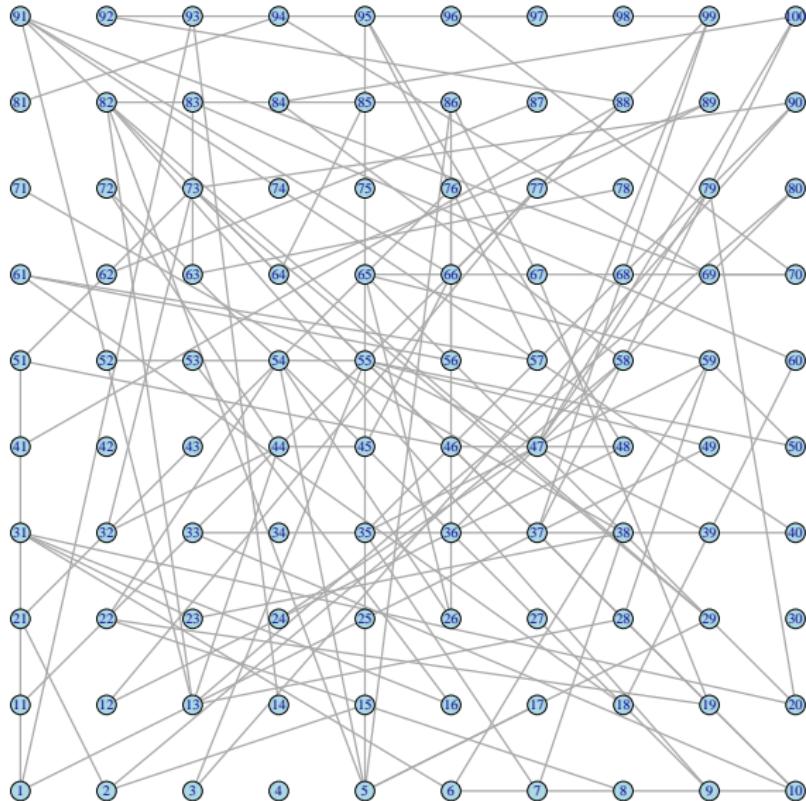
Network visualisation [recap]

Layout: Circle



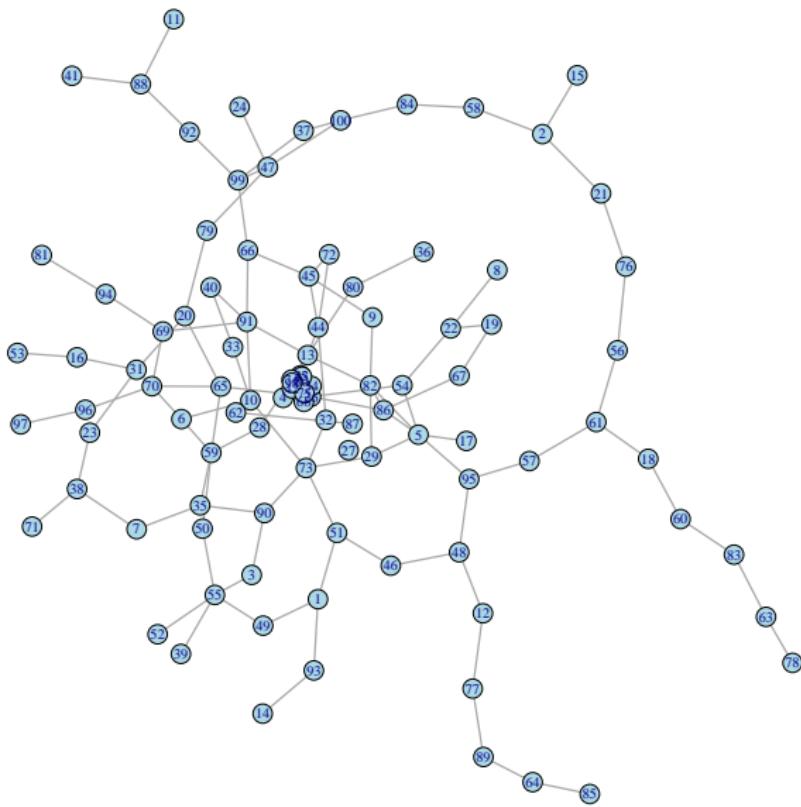
Network visualisation [recap]

Layout: Grid



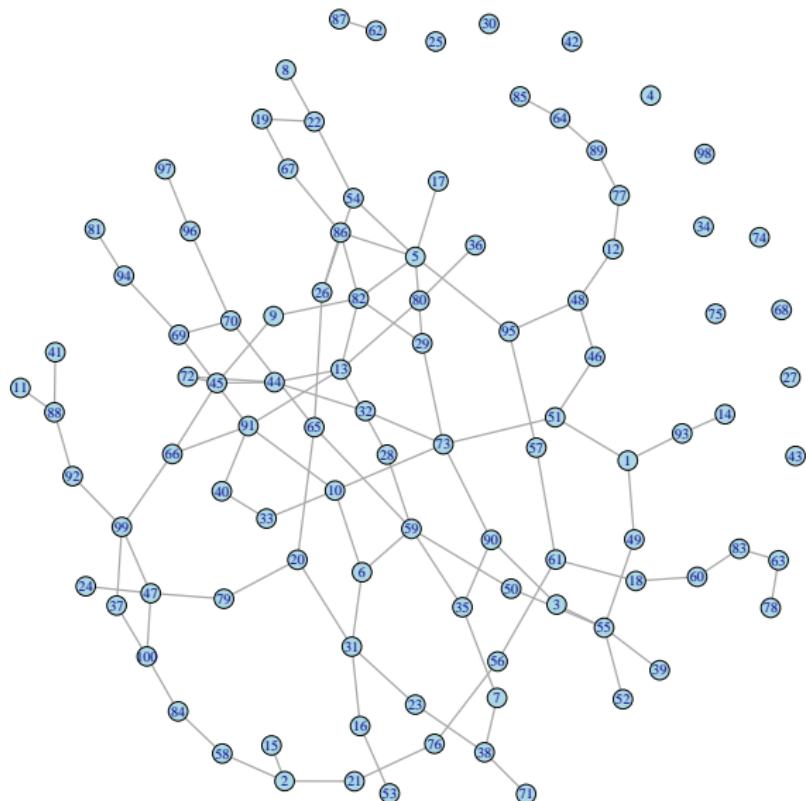
Network visualisation [recap]

Layout: Kamada-Kawai



Network visualisation [recap]

Layout: Fruchterman and Reingold



Network visualisation [recap]

Let's practice this in RStudio (open S7_Script.R)

Introduction to Gephi



<https://gephi.org/>

- Developed by the **Gephi Consortium** (a not-for-profit organisation)
- Runs on **Windows, Linux, Mac OSX**
- Relatively **user-friendly**
- Can produce excellent **network visualizations**
- Can read several **file formats**
- Capable of analysing **relatively large networks**
- Gephi plugins (<https://gephi.org/plugins>)
- Operation cannot be **automated**

Introduction to Gephi

Interfaces

- **Overview**

- ▶ to generate network layouts
- ▶ to colour nodes/edge
- ▶ to calculate network measures

- **Data laboratory**

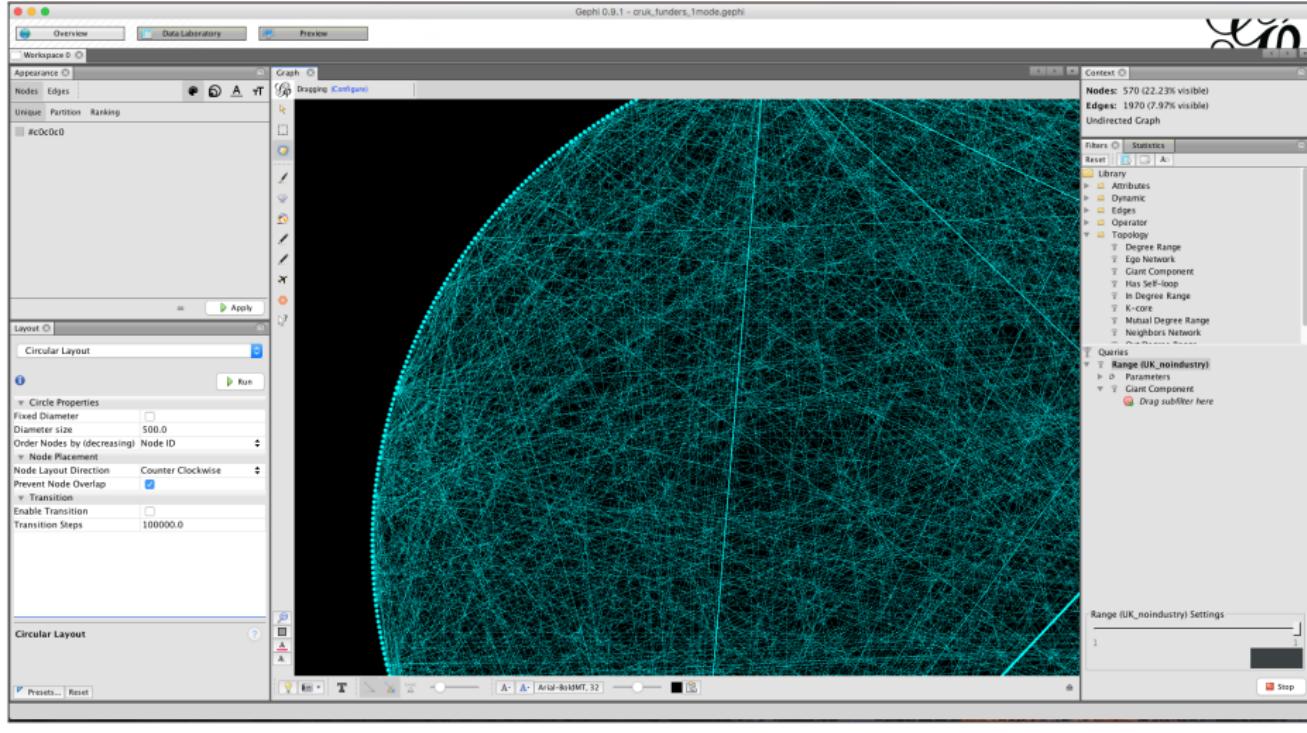
- ▶ to upload network data
- ▶ to manipulate network data
- ▶ to export network data

- **Preview**

- ▶ to preview network images
- ▶ to generate network images (e.g. pdf, jpeg)

Introduction to Gephi

Overview interface



Introduction to Gephi

Data laboratory interface



Gephi 0.9.1 - crux_funders, 1mode.gexf

Workspace D Data Laboratory Preview

Nodes Edges Configuration Add node Add edge Search/Replace Import Spreadsheet Export table More actions

ID	Name	Interval	LOGPS	Component ID
2574	Aberdeen Medico-Chirurgical Society		1	0
2575	Aberdeen Proton Therapy Facility		1	0
2576	Above & Below Beyone		2	0
2581	Academy of Medical Sciences		2	0
2585	ACT		1	0
2588	Action Against Bladder Cancer		1	0
2589	Action Cancer Northern Ireland		11	0
2592	Addenbrookes Charitable Trust		16	0
2597	Against West Midlands		1	0
2602	Against Breast Cancer		2	0
2603	Age UK		2	0
2612	AIDS Research Trust		1	0
2613	Air and Ground		1	0
2618	AKU Society		1	0
2619	Alan Meirement Memorial Fund		3	0
2628	Als Dream		4	0
2629	Alan J Lerner Fund		2	0
2637	Alpha Omega Research Trust		2	0
2672	Anglo-European College of Chiropractic		1	0
2675	Annabel McElroy Childrens Cancer Fund		1	0
2679	Antenatal Results and Choices		1	0
2685	Arthritis Foundation		5	0
2697	Arthritis Research UK.		8	0
2708	Association for International Cancer Research		54	0
2712	Association of Anaesthetists of Great Britain and Ireland		1	0
2713	Association of Clinical Biochemists		3	0
2714	Association of Clinical Immunologists		3	0
2715	Association of Medical Microbiologists		2	0
2716	Association of Palliative Medicine		2	0
2727	Association of British Neurologists		1	0
2731	Asthma UK		2	0
2733	Axata Telangiectasia Society		2	0
2736	Audrey Callaghan Fellowship		1	0
2765	Babraham Institute		1	0
2773	Bardens Research and Education Trust		2	0
2774	Barry Research Fund		2	0
2775	Barts and The London Charitable Foundation		11	0
2785	Beaton Institute for Cancer Research		2	0
2786	Beaumont Oncology Centre		1	0
2794	Belfast City Hospital		1	0
2795	Belfast City Hospital Charitable Trust Funds		1	0
2796	Belfast Gynaecological Cancer Network		1	0
2800	BeHouse Foundation		1	0

Add column Merge columns Delete column Clear column Copy data to other column Fill column with a value Duplicate column Create a boolean column from regex match Create column with list of regex matching groups Negate boolean values Convert column to dynamic

Navigation icons: back, forward, search, etc.

Introduction to Gephi

Preview interface



Gephi 0.9.1 - cruk_funders.tmode.gexf

Overview Data Laboratory Preview

Presets Default

Settings Manage renderers

Nodes

- Border Width: 1.0
- Border Color: custom [0,0,0]
- Opacity: 100.0

Node Labels

- Show Labels:
- Font: Arial 12 Plain
- Proportional size:
- Color: custom [255,255,255]
- Shorten label:
- Max characters: 30
- Outline size: 0.0
- Outline color: parent
- Outline opacity: 80.0

Box

- Box color: parent
- Box opacity: 100.0

Edges

- Show Edges:
- Thickness: 1.0
- Rescale weight:
- Color: mixed
- Opacity: 70.0
- Curved:
- Radius: 0.0

Edge Arrows

- Size: 3.0

Edge Labels

- Show Labels:
- Font: Arial 10 Plain
- Color: custom [255,255,255]
- Shorten label:
- Max characters: 30
- Outline size: 0.0
- Outline color: parent
- Outline opacity: 80.0

Preview ratio: 100%

Export: SVG/PDF/PNG Refresh

Background Reset zoom - +

Introduction to Gephi

Ready?

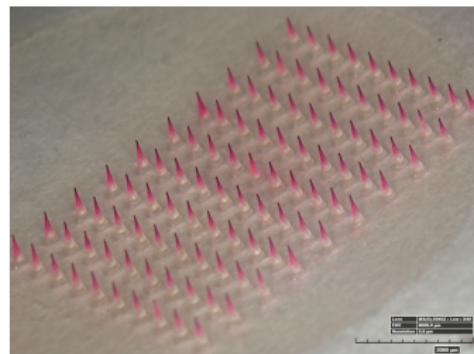
- ① Go to <https://gephi.org/>
- ② Download Gephi
- ③ Install Gephi*

*If you are using a computer on campus, install Gephi on your 'N' drive

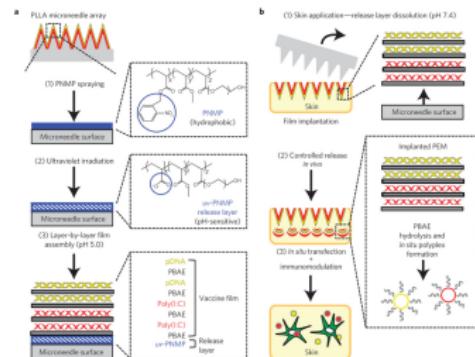
Introduction to Gephi

Exercise

Microneedles: needles the size of which is on the *micrometer* length scale. Expected impact on vaccines, drug delivery, and reduction of biohazard waste



Source: www.news.gatech.edu



Source: DeMuth PC et al. (2013). Polymer multilayer tattooing for enhanced DNA vaccination. *Nature Materials*, 12(4), 367-276

Introduction to Gephi

Exercise

Objectives of the analysis

- To map collaboration activity
- To characterise the inter-organisational network
- To identify key organisations in this network

Introduction to Gephi

Exercise

1. Data collection

- Search for expert-defined keywords in publication titles and abstracts
 - ▶ *microneedle**
 - ▶ *micro-needle**
 - ▶ *micropjection patch**
 - ▶ *micro-projection patch**
 - ▶ ...
- 1,090 publications in the 2010-2014 period
- 887 organizations (disambiguation of names)
- Classification of the organisations by type
 - ▶ Research and Higher Education (RHE)
 - ▶ Healthcare Provider (HCP)
 - ▶ Government (GOV)
 - ▶ Research Institute (RIN)
 - ▶ Industry (IND)
 - ▶ Non-Government Organization (NGO)
- **Assumption:** Co-authorship = Collaboration (!)

Introduction to Gephi

Exercise

2. Data preparation

- Data were pre-processed
- Gephi can import **Excel files**
- We need to create two files or sheets
 - ▶ Nodes table
 - ▶ Edges table
- Format these 'tables' as in the file
"microneedles_2010_14_for_gephi.xlsx"

	A	B	C	D
1	Id	Label	Country	Type
2	1	3M Drug Delivery Systems Div (Singapore)	Singapore	IND
3	2	3M Drug Delivery Systems Div (United States)	United States	IND
4	3	Aalborg Univ (Denmark)	Denmark	RHE
5	4	ABS Global (United States)	United States	IND
6	5	Acad Pharma Inc (United States)	United States	IND

Edges table

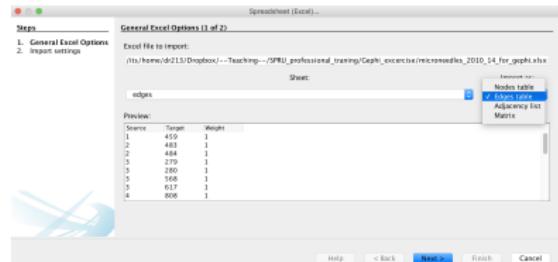
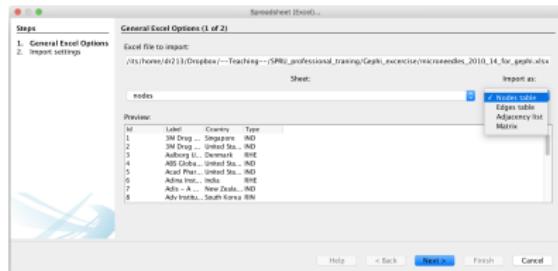
	A	B	C	
1	Source	Target	Weight	
2		1	459	1
3		2	483	1
4		2	484	1
5		3	279	1
6		3	280	1

Introduction to Gephi

Exercise

3. Import data into Gephi

- ① Create a new project in Gephi
- ② Select *Data Laboratory*
- ③ Select *Import Spreadsheet*
 - ▶ We first import nodes, then edges
 - ▶ Remember to import the network as undirected
 - ▶ Make sure you tick the *Append to existing workspace* option
- ④ Save the project

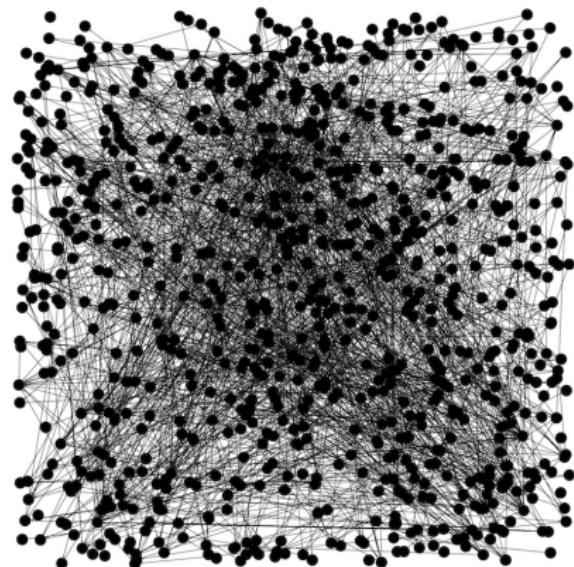


Introduction to Gephi

Exercise

4. Improve the visualisation

- Colour nodes by type
 - ① Go to *Appearance* in *Overview*
 - ② Select *Nodes, Colors, and Partition*
 - ③ Select the attribute type and *Apply*
- Improve the layout
 - ① Go to the *Layout* tab
 - ② Select a layout algorithm (*Yifan Hu*)
 - ③ Try different layout algorithms
- Exclude isolates and small components
 - ① Go to the *Filters* tab
 - ② Expand the folder *Topology*
 - ③ Drag and drop *Giant Component* into *Queries and Filter*
- Increase node size by degree
 - ① Go to *Appearance*
 - ② Select *Nodes, Size, and Ranking*
 - ③ Select the *Degree* attribute (range from 10 to 50) and *Apply*

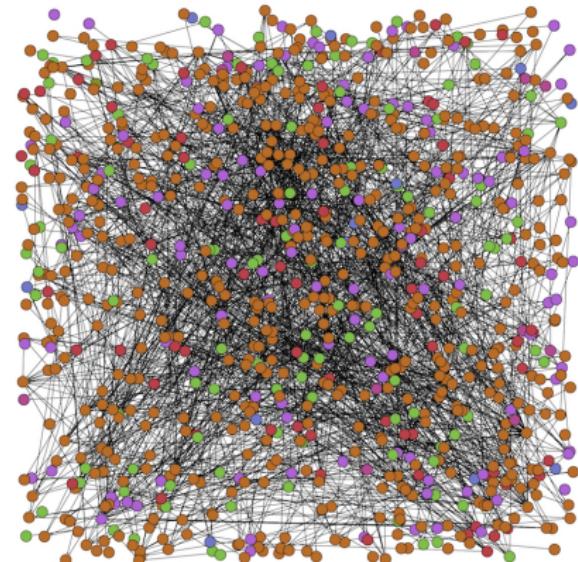


Introduction to Gephi

Exercise

4. Improve the visualisation

- Colour nodes by type
 - ① Go to *Appearance* in *Overview*
 - ② Select *Nodes, Colors, and Partition*
 - ③ Select the attribute type and *Apply*
- Improve the layout
 - ① Go to the *Layout* tab
 - ② Select a layout algorithm (*Yifan Hu*)
 - ③ Try different layout algorithms
- Exclude isolates and small components
 - ① Go to the *Filters* tab
 - ② Expand the folder *Topology*
 - ③ Drag and drop *Giant Component* into *Queries and Filter*
- Increase node size by degree
 - ① Go to *Appearance*
 - ② Select *Nodes, Size, and Ranking*
 - ③ Select the *Degree* attribute (range from 10 to 50) and *Apply*

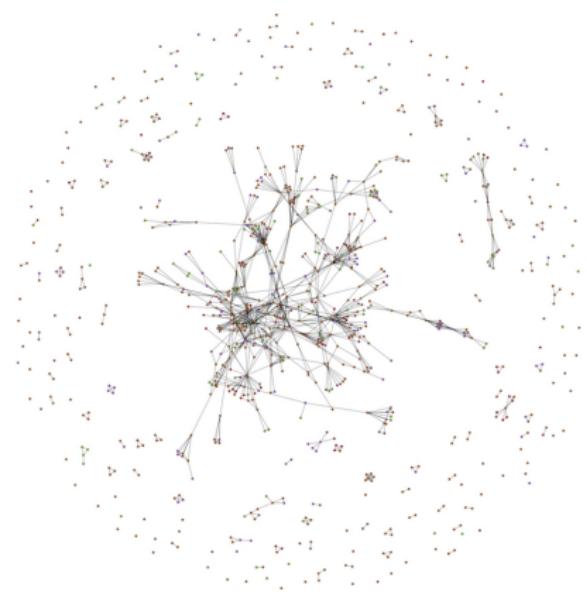


Introduction to Gephi

Exercise

4. Improve the visualisation

- Colour nodes by type
 - ① Go to *Appearance* in *Overview*
 - ② Select *Nodes, Colors, and Partition*
 - ③ Select the attribute type and *Apply*
- Improve the layout
 - ① Go to the *Layout* tab
 - ② Select a layout algorithm (*Yifan Hu*)
 - ③ Try different layout algorithms
- Exclude isolates and small components
 - ① Go to the *Filters* tab
 - ② Expand the folder *Topology*
 - ③ Drag and drop *Giant Component* into *Queries and Filter*
- Increase node size by degree
 - ① Go to *Appearance*
 - ② Select *Nodes, Size, and Ranking*
 - ③ Select the *Degree* attribute (range from 10 to 50) and *Apply*

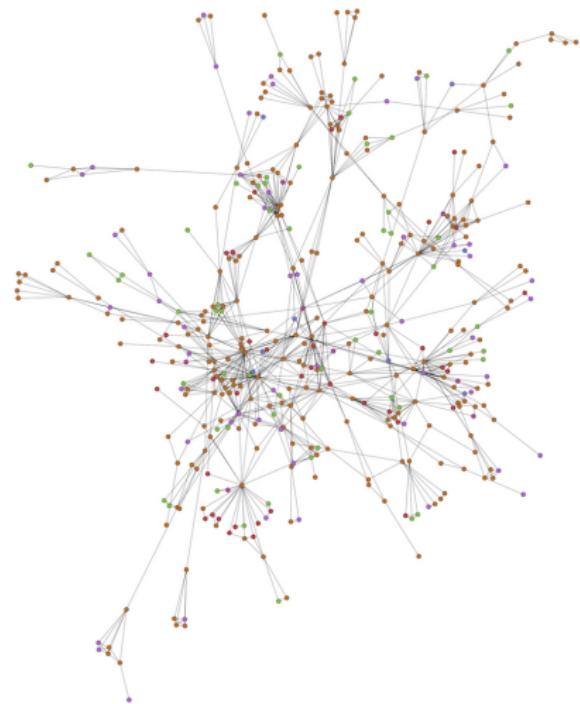


Introduction to Gephi

Exercise

4. Improve the visualisation

- Colour nodes by type
 - ① Go to *Appearance* in *Overview*
 - ② Select *Nodes, Colors, and Partition*
 - ③ Select the attribute type and *Apply*
- Improve the layout
 - ① Go to the *Layout* tab
 - ② Select a layout algorithm (*Yifan Hu*)
 - ③ Try different layout algorithms
- Exclude isolates and small components
 - ① Go to the *Filters* tab
 - ② Expand the folder *Topology*
 - ③ Drag and drop *Giant Component* into *Queries and Filter*
- Increase node size by degree
 - ① Go to *Appearance*
 - ② Select *Nodes, Size, and Ranking*
 - ③ Select the *Degree* attribute (range from 10 to 50) and *Apply*

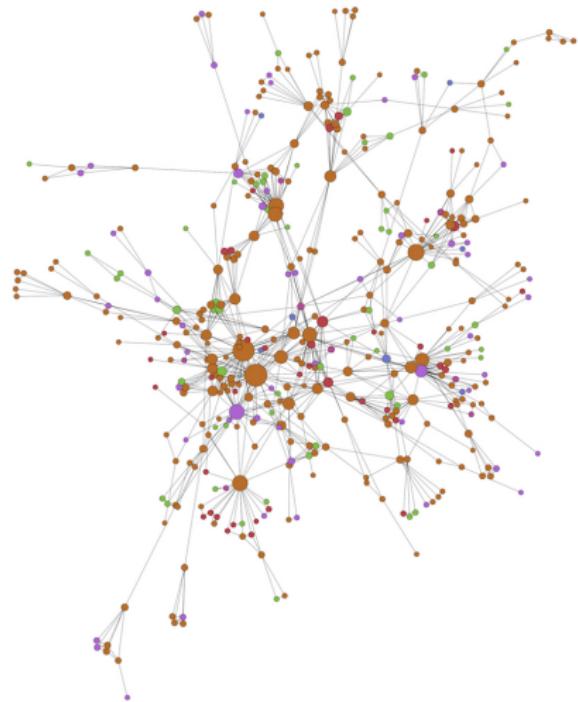


Introduction to Gephi

Exercise

4. Improve the visualisation

- Colour nodes by type
 - ① Go to *Appearance* in *Overview*
 - ② Select *Nodes, Colors, and Partition*
 - ③ Select the attribute type and *Apply*
- Improve the layout
 - ① Go to the *Layout* tab
 - ② Select a layout algorithm (*Yifan Hu*)
 - ③ Try different layout algorithms
- Exclude isolates and small components
 - ① Go to the *Filters* tab
 - ② Expand the folder *Topology*
 - ③ Drag and drop *Giant Component* into *Queries and Filter*
- Increase node size by degree
 - ① Go to *Appearance*
 - ② Select *Nodes, Size, and Ranking*
 - ③ Select the *Degree* attribute (range from 10 to 50) and *Apply*

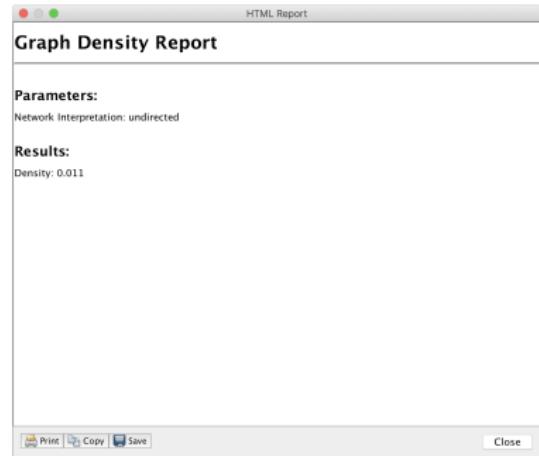


Introduction to Gephi

Exercise

5. Characterise the network

- ① We can calculate the density
- ② Go to *Statistics*
- ③ Run *Graph Density*
- ④ Giant component's density = 0.011
- ⑤ Other measures: Diameter, Components, Modularity, etc.



Introduction to Gephi

Exercise

6. Identify key organisations

- ① We can calculate centrality measures
- ② Run *Average Degree*
- ③ Run *Network Diameter*
- ④ Closeness and betweenness are added as attributes (see *Data Laboratory*)

Organisation	Degree
Georgia Inst of Tech (United States)	41
Emory Univ (United States)	39
Univ of Queensland (Australia)	27
Yonsei Univ, Seoul (South Korea)	25
Korea Adv Inst of Sci and Tech (KAIST) (South Korea)	24
North Carolina State Univ (United States)	24
Harvard Univ (United States)	23
Univ of North Carolina at Chapel Hill (United States)	23
Queen's Univ, Belfast (United Kingdom)	21
Massachusetts Inst of Tech (United States)	20
...	

Introduction to Gephi

Exercise

6. Identify key organisations

- ① We can calculate centrality measures
- ② Run *Average Degree*
- ③ Run *Network Diameter*
- ④ Closeness and betweenness are added as attributes (see *Data Laboratory*)

Organisation	Closeness
Georgia Inst of Tech (United States)	0.31
Emory Univ (United States)	0.31
Massachusetts Inst of Tech (United States)	0.30
Korea Adv Inst of Sci and Tech (KAIST) (South Korea)	0.29
Harvard Univ (United States)	0.29
Seoul Nat Univ (South Korea)	0.29
Cardiff Univ (United Kingdom)	0.28
Sungkyunkwan Univ (South Korea)	0.28
Kangbuk Samsung Hosp (South Korea)	0.28
Univ of Oxford (United Kingdom)	0.27
...	

Introduction to Gephi

Exercise

6. Identify key organisations

- ① We can calculate centrality measures
- ② Run *Average Degree*
- ③ Run *Network Diameter*
- ④ Closeness and betweenness are added as attributes (see *Data Laboratory*)

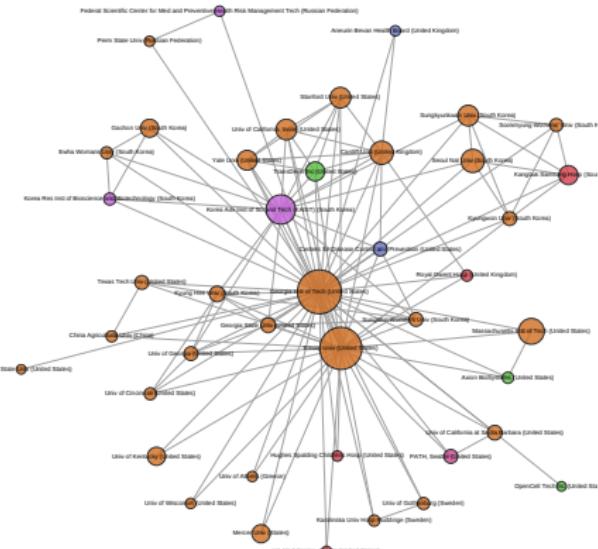
Organisation	Betweenness
Georgia Inst of Tech (United States)	19108
Osaka Univ (Japan)	16665
Emory Univ (United States)	16378
Cardiff Univ (United Kingdom)	14442
Harvard Univ (United States)	13327
Univ of Queensland (Australia)	13072
Massachusetts General Hosp (United States)	12001
Korea Adv Inst of Sci and Tech (KAIST) (South Korea)	11620
Seoul Nat Univ (South Korea)	10648
Massachusetts Inst of Tech (United States)	10631
...	

Introduction to Gephi

Exercise

7. Visualising ego networks

- ① We can visualise the ego network of the Georgia Inst of Tech
- ② Go to *Filters*
- ③ Expand the folder *Topology*
- ④ Drag an drop *Ego Network* into *Queries*
- ⑤ Select *Node Id* and *Filter*



Next time ...

Next time ...

- **Lecture: Network models**

- ▶ Mathematical models of network analysis
- ▶ Overview of statistical models of network analysis

- **Seminar: Network models**

- ▶ Network mathematical models in igraph
- ▶ Gephi exercise