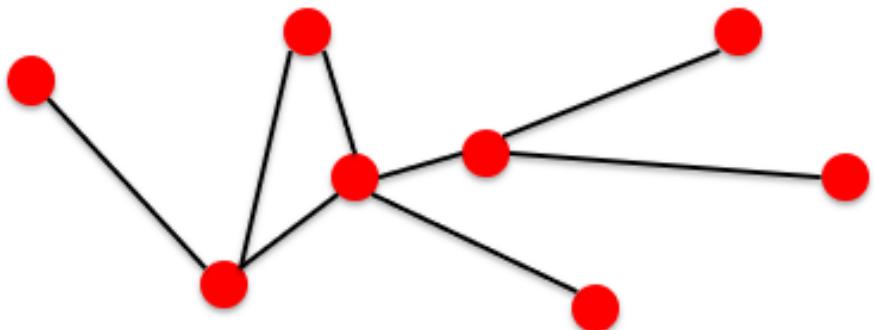


Network Analysis and Visualization

Veera Muangsin

Network represents relationships



- **components:** nodes, vertices
- **interactions:** links, edges
- **system:** network, graph

N
L
(N,L)

Examples

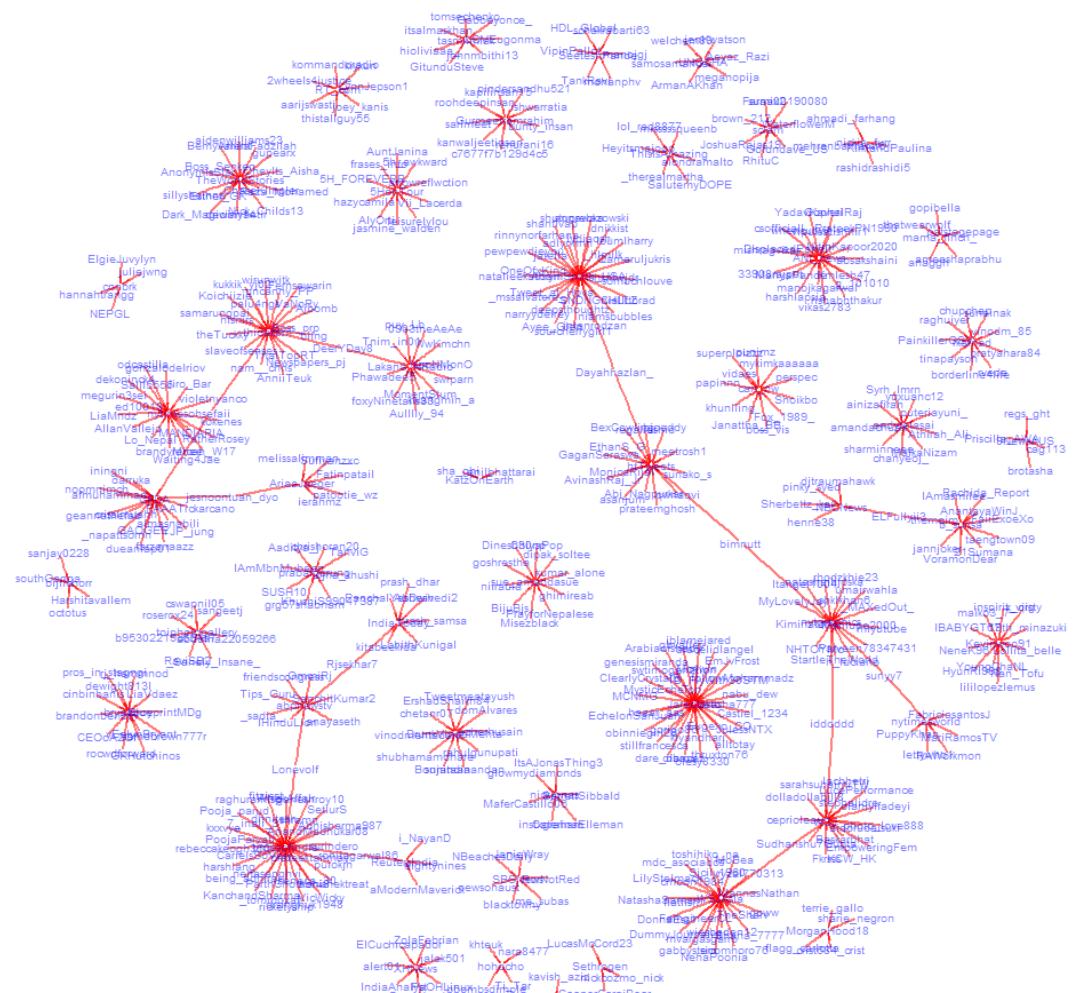
- Contacts (telephone, email)
- Social networks (friend, follow, share)
- Reference (citation, hyperlinks)
- Collocation/co-occurrence (words, events)
- Connections (roads, computer network)

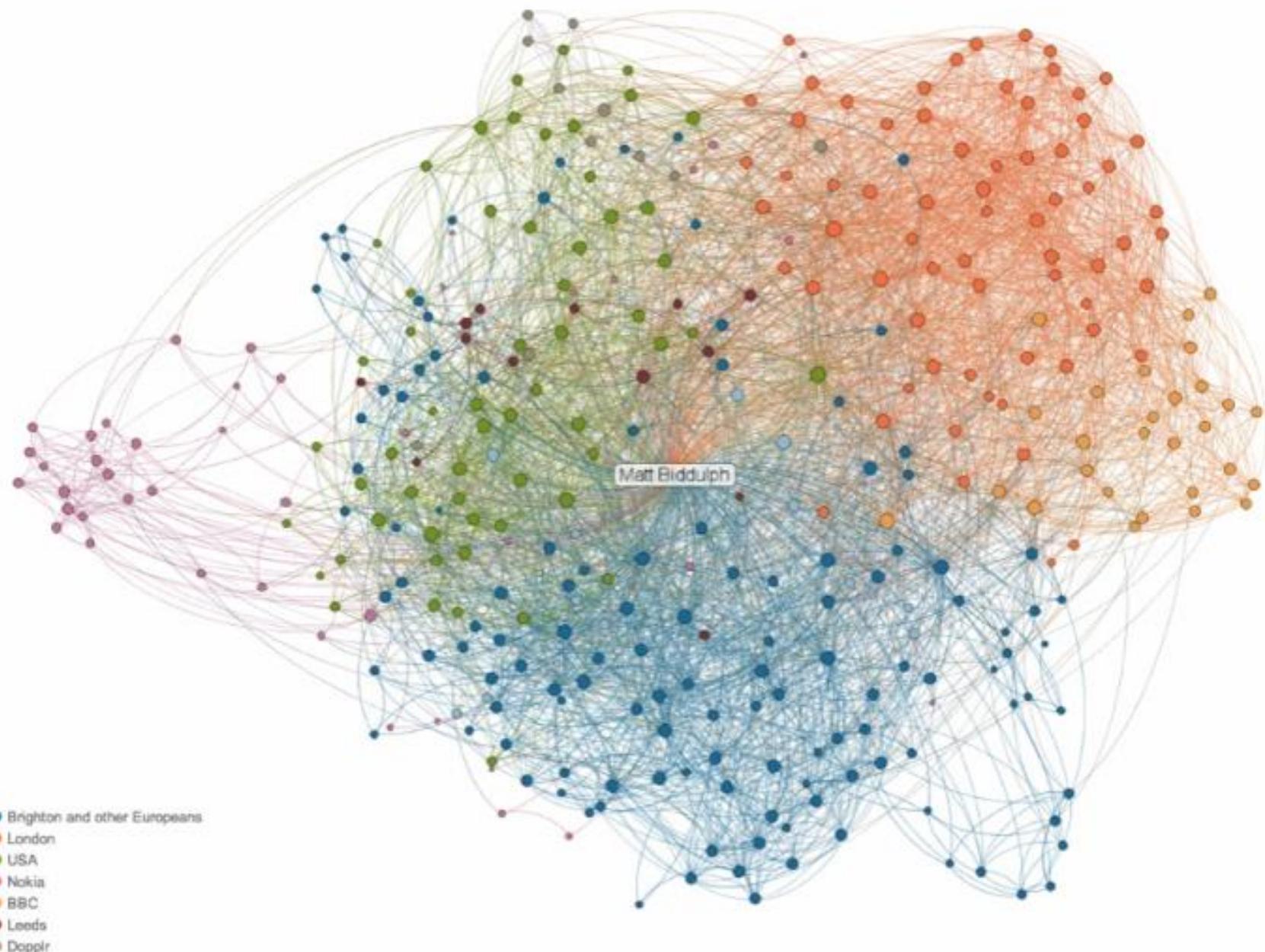
Applications

- Social Network Analysis
- Bibliographic Analysis
- Transportation/communication

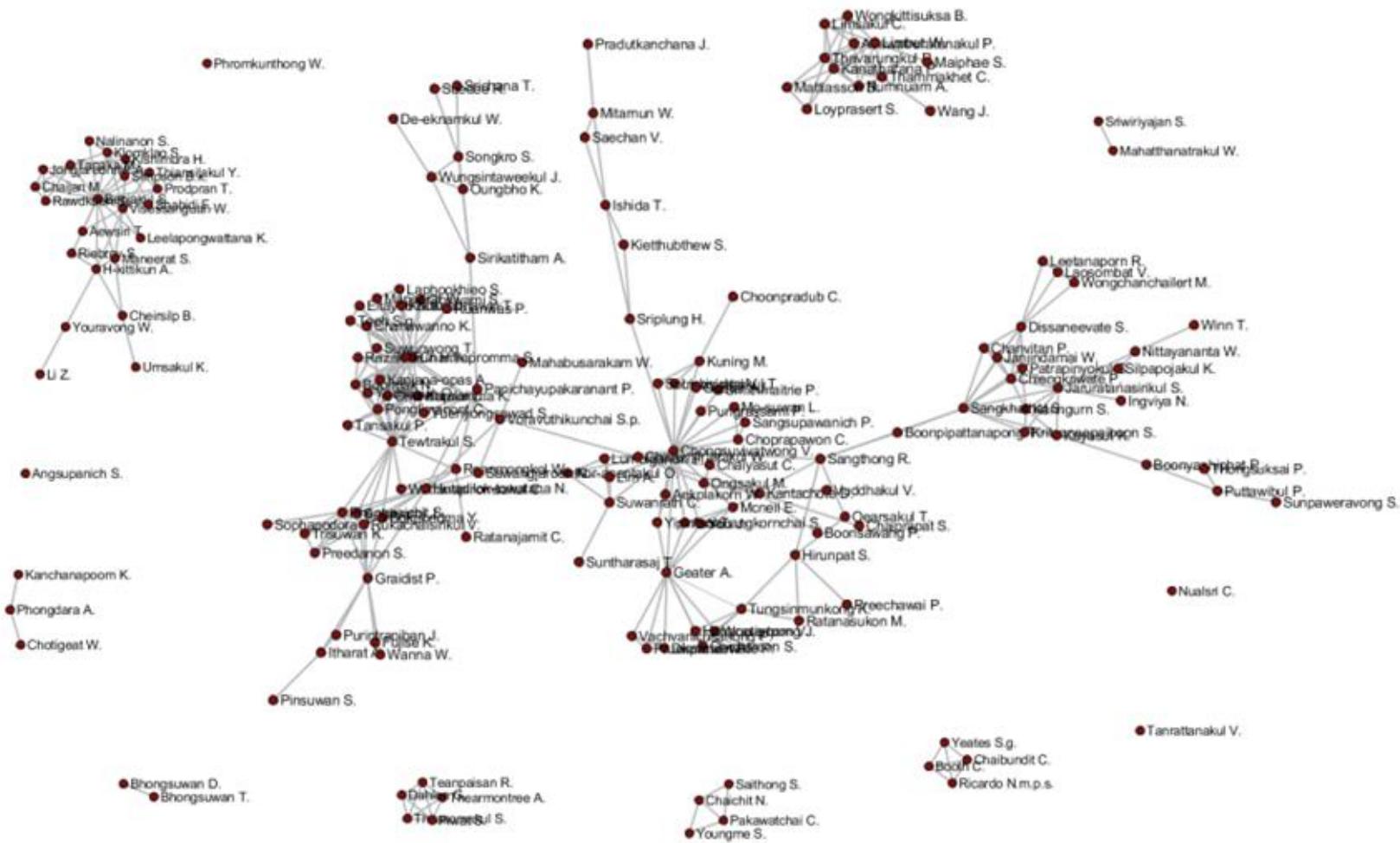
Retweet Network

Degree : ດັບຮອນເກືອນຫຼຸດ





Co-authorship Network



Ingredient Complement Network (word collocation)

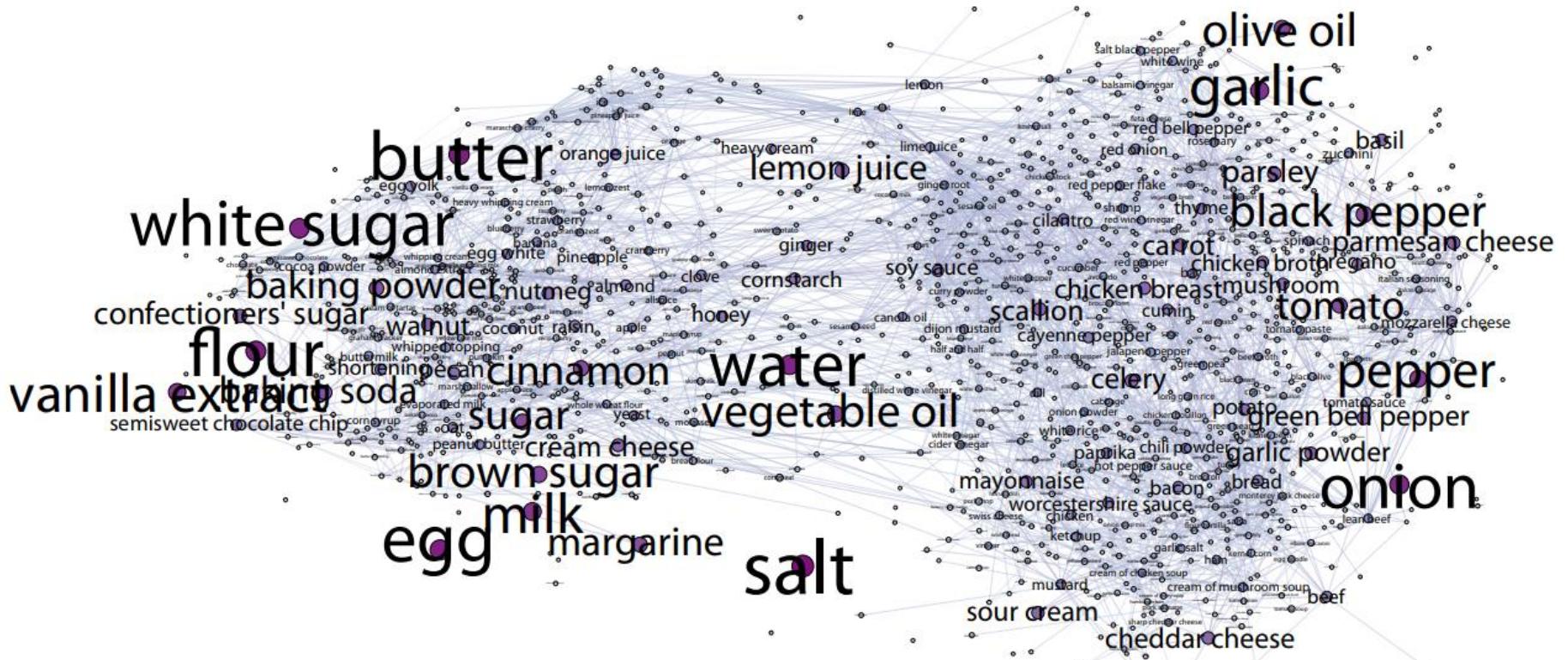
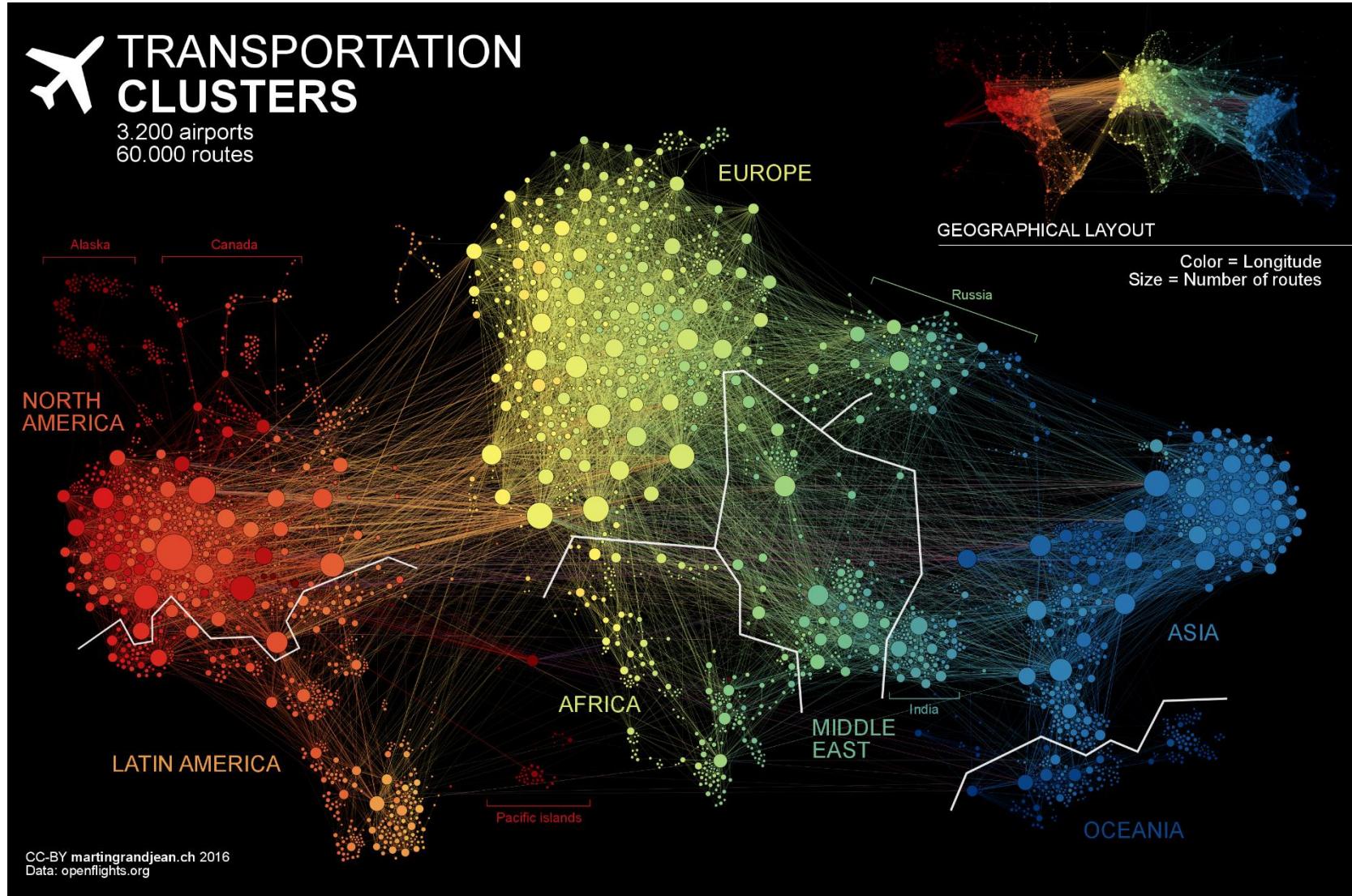
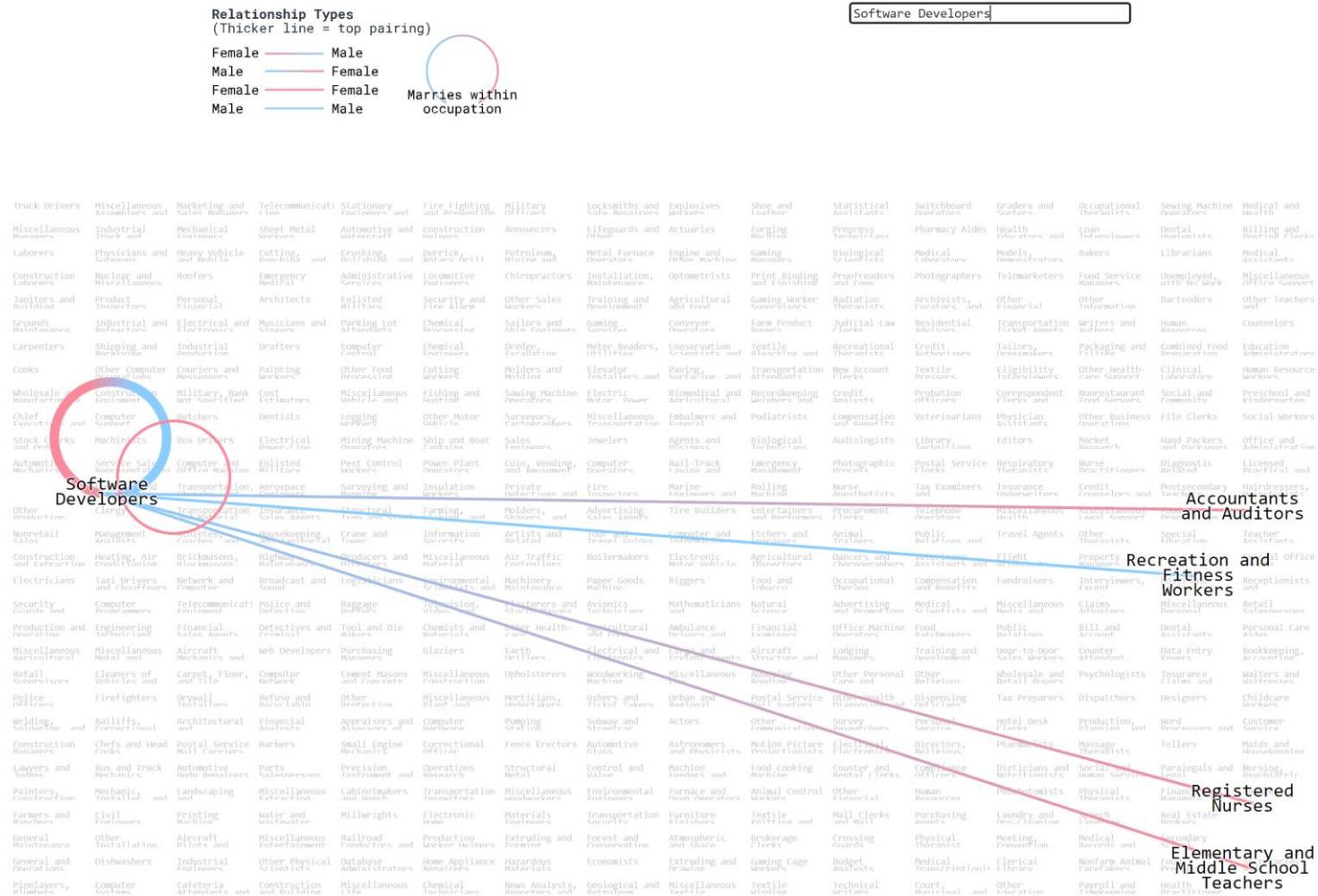


Figure 2: Ingredient complement network. Two ingredients share an edge if they occur together more than would be expected by chance and if their pointwise mutual information exceeds a threshold.

Air Traffic Network



Who marries whom (interactive)



<https://www.bloomberg.com/graphics/2016-who-marries-whom/>

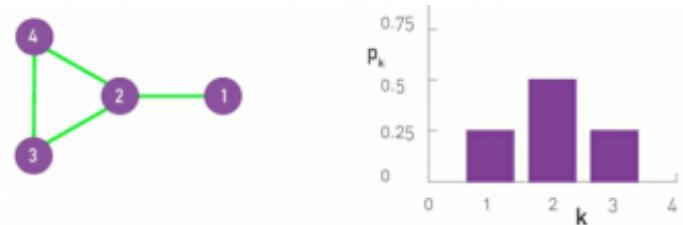
Network Models

Basic network models

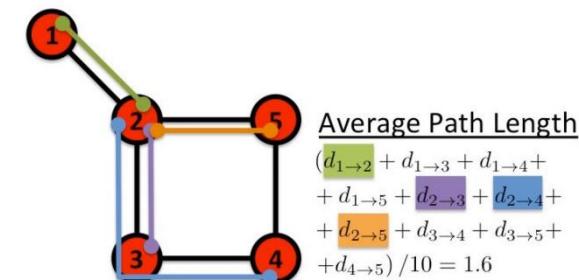
- Random network
- Small-world network
- Scale-free network

Key Network Features

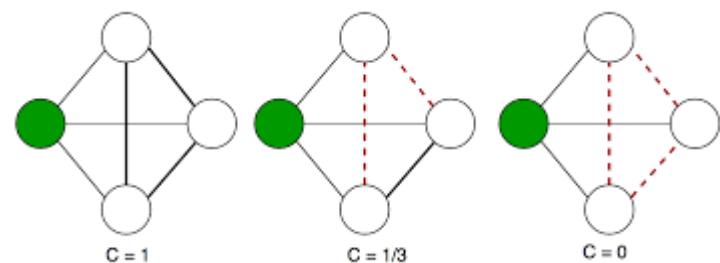
- Degree Distribution *ទំនាក់ទំនង*
 - Distribution function of node degrees
 - $P(k)$ = fraction of nodes with degree k



- Average Path Length
 - Average number of hops between any two nodes



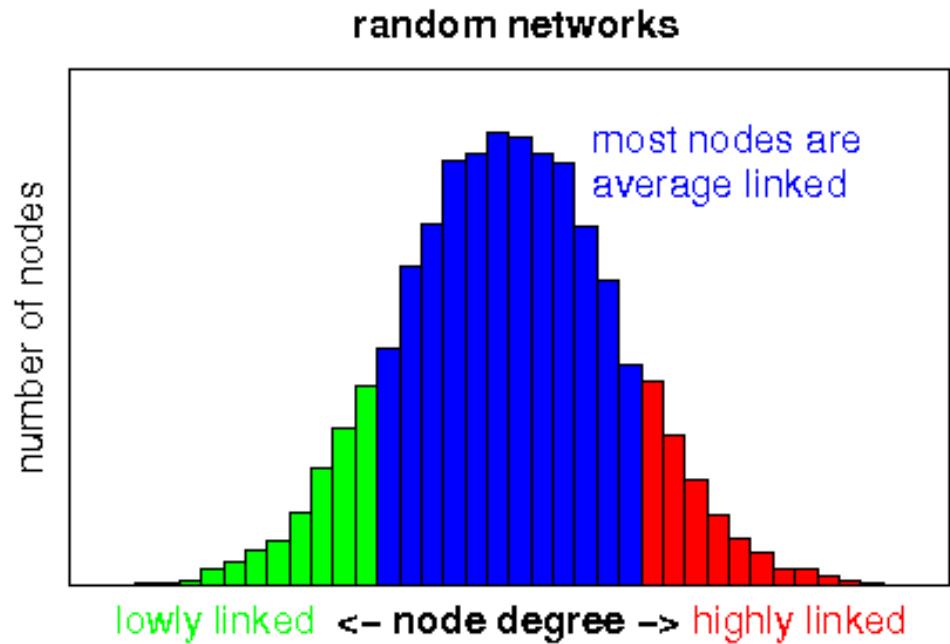
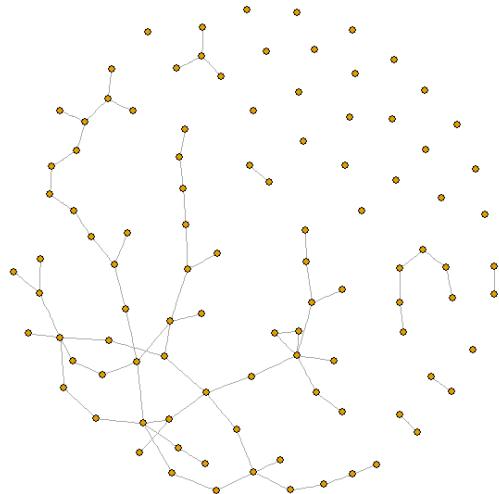
- Clustering coefficient
 - Average fraction of pairs of neighbors of a node that are also neighbors of each other.



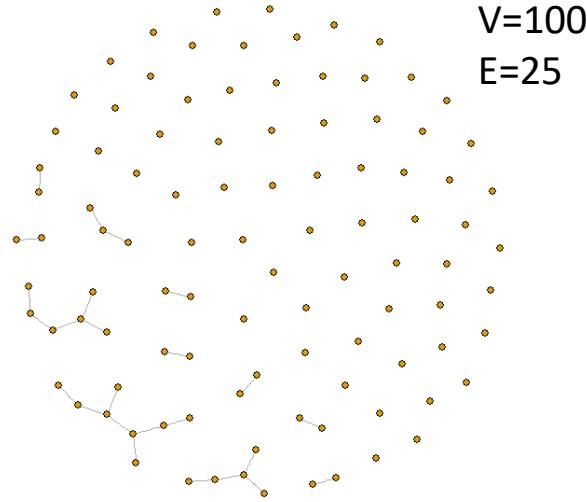
$$C = \frac{\text{no. of triangles}}{\text{no. of all possible triangles}}$$

Random Network

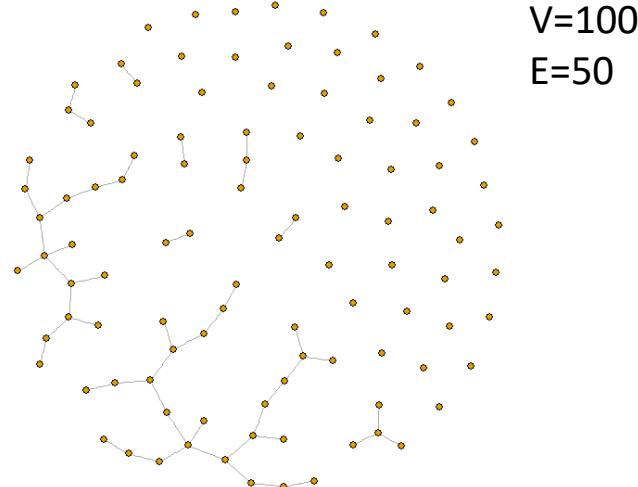
- A random network is obtained by starting with a set of n isolated vertices and adding successive edges between them at random.
- A giant component (largest connected component) will emerge.
- Degree distribution is normal distribution.
- Random networks may not represent real-world networks but offer base-line comparison to other models.
- P. Erdős and A. Rényi model



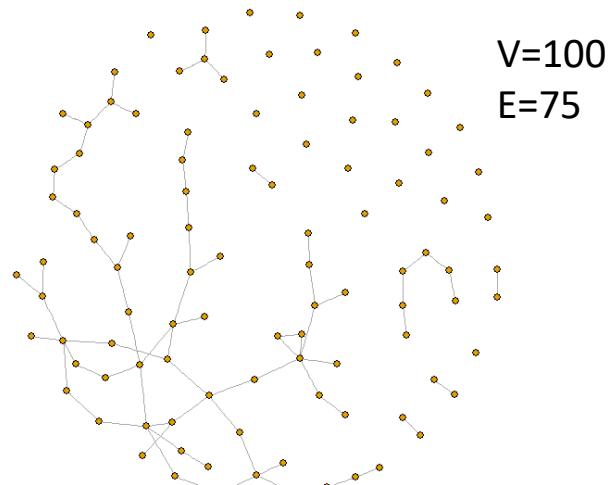
Random network



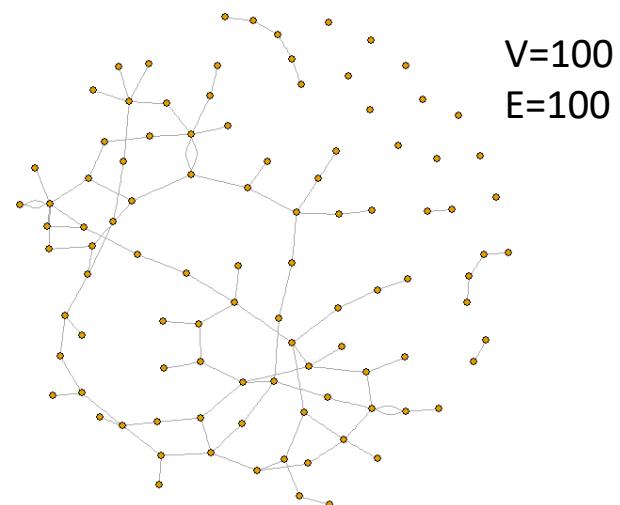
$V=100$
 $E=25$



$V=100$
 $E=50$



$V=100$
 $E=75$



$V=100$
 $E=100$

Small-world Network

node laying
edge laying

- Network with high clustering coefficient and short average path length
- Most nodes are connected to a few nearby neighbors but some nodes also have connections to far away nodes.
- That shortens the average path length and therefore most nodes can be reached from every other node by a small number of hops.
- Real-world example of small-world network
 - Social connection
 - Email, telephone call-graph
- Watts-Strogatz model

Rewiring of Links

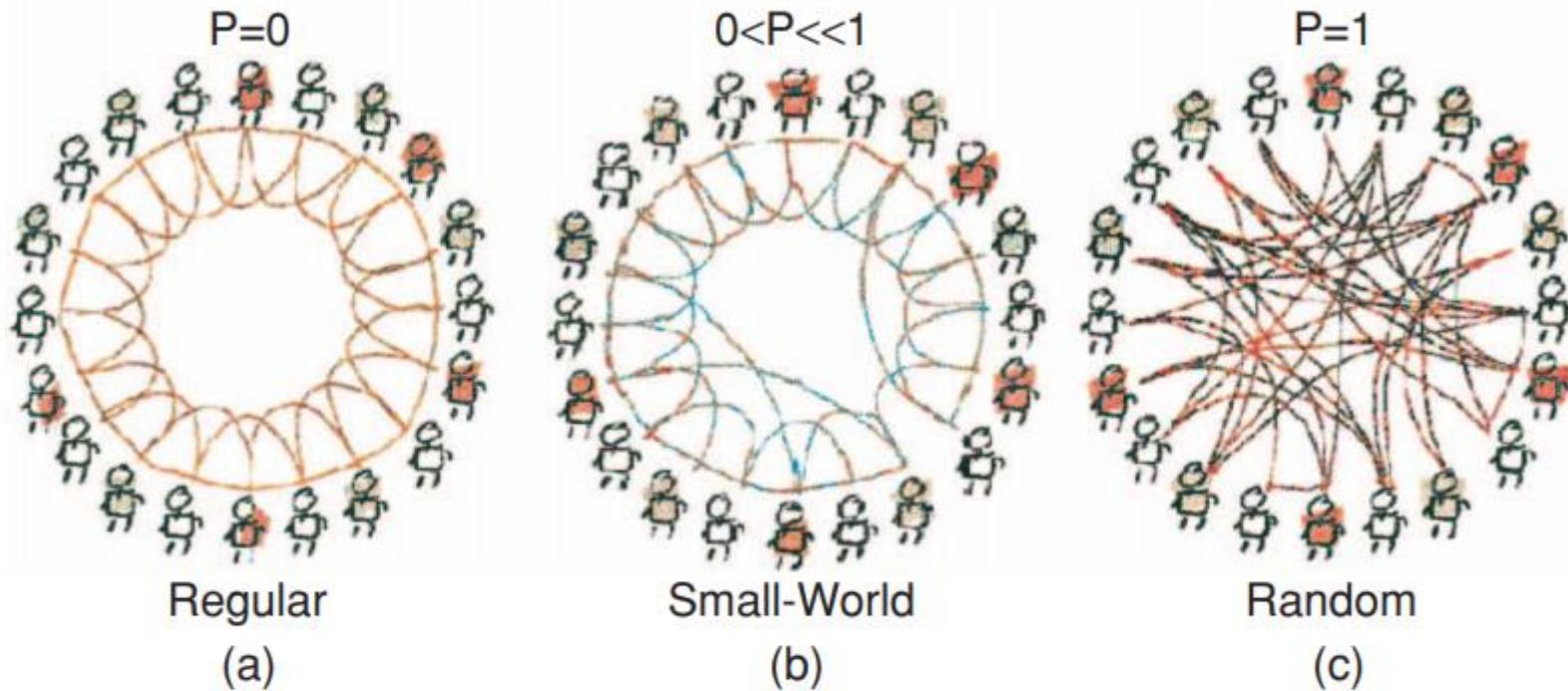
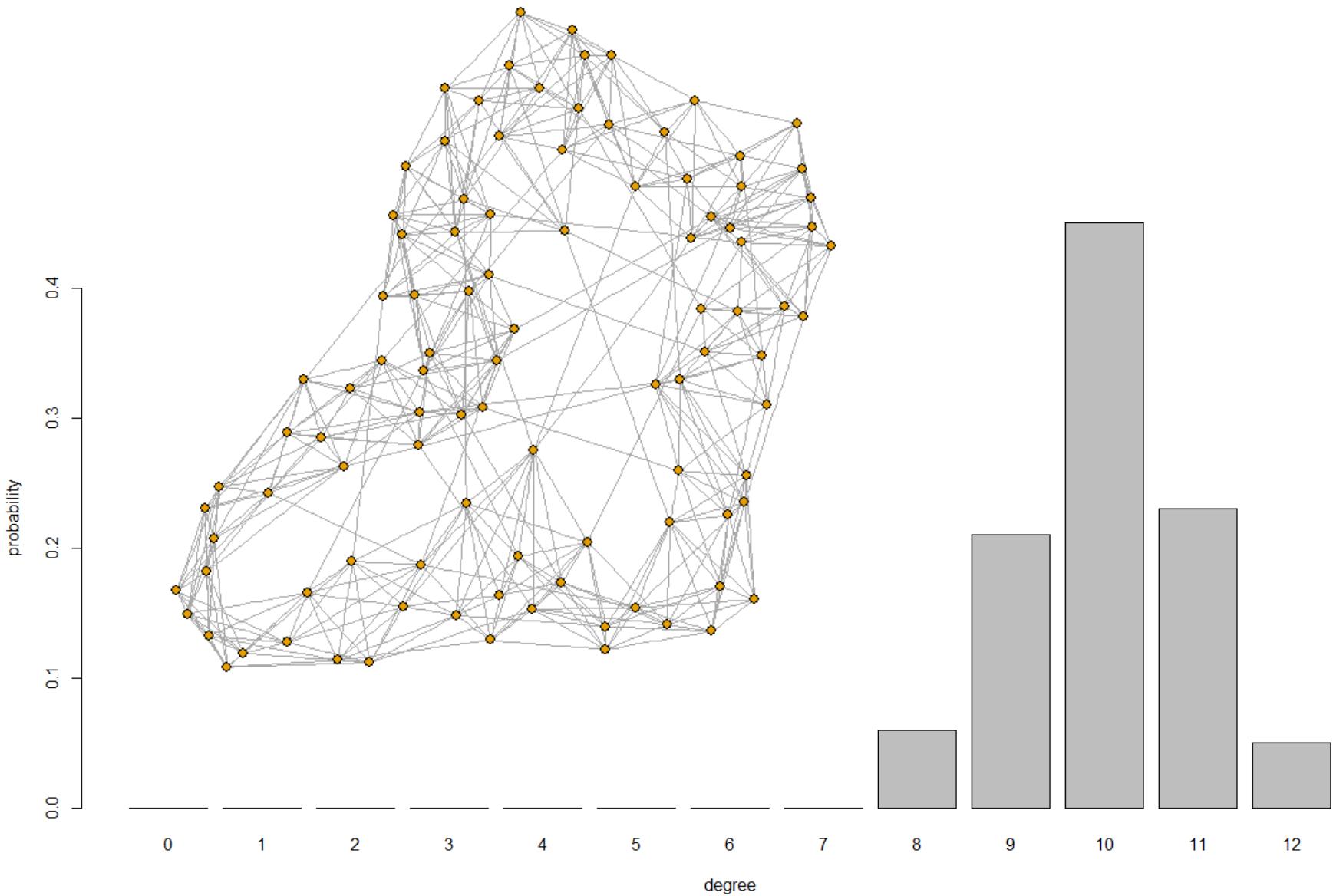


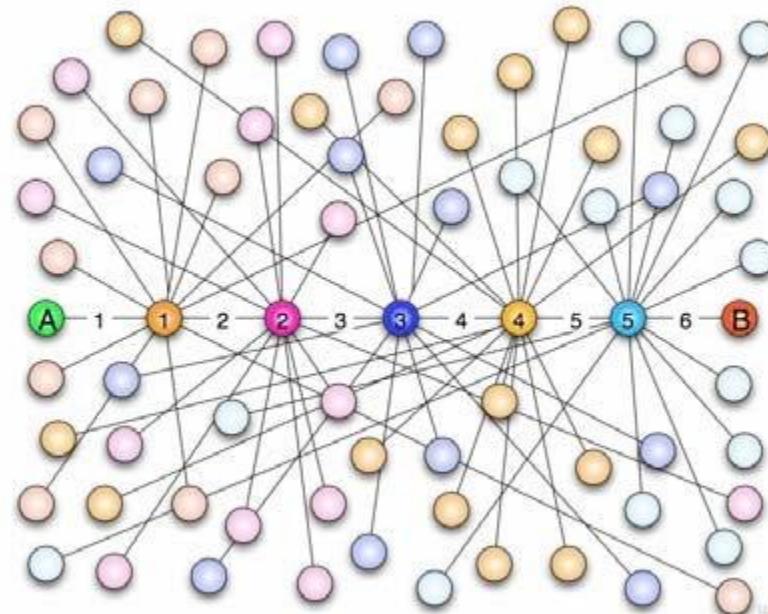
Figure 7. (a) In this completely regular friendship network, people are friends with **only their 4 nearest neighbors**. The network is highly cliquish, and any 2 people are on average many degrees apart. (b) In this small-world network, people still know **4 others on average**, but a few have **distant friends**. The network is still **highly cliquish**, but the **average degree of separation is small**. (c) In this random network, everyone **still knows 4 others on average**, but **friends are scattered**: few people have many friends in common, and pairs are on average only a **few degrees apart**.

A Small-World Network



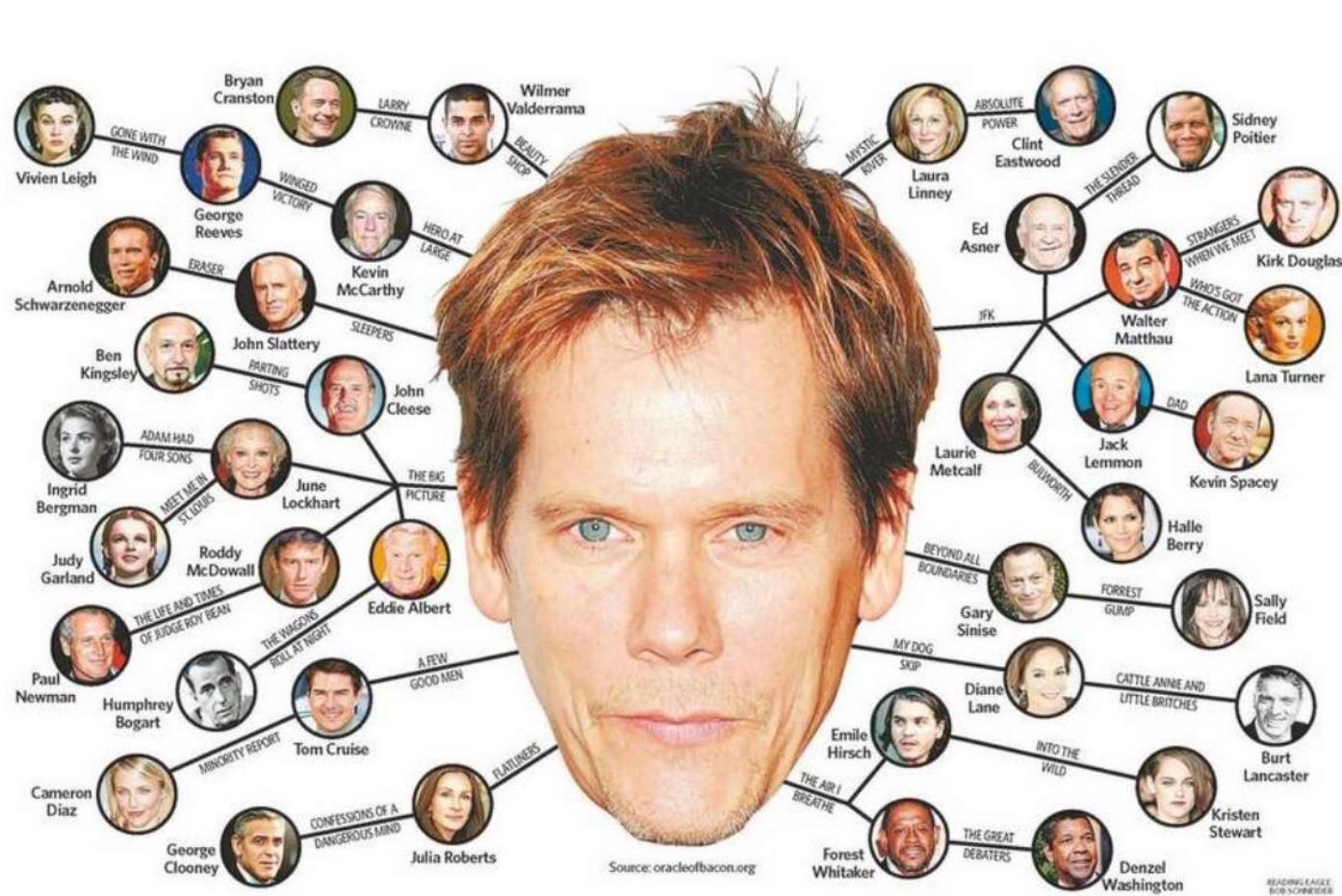
Six degrees of separation

- A theory that anyone can connect to anyone else in the world via at most six steps of social connection.



Six degree of Kevin Bacon

- A prediction that any actor can be linked through their film roles to Bacon within six steps



<https://www.youtube.com/watch?v=n9u-TITxwoM>

Scale-Free Network

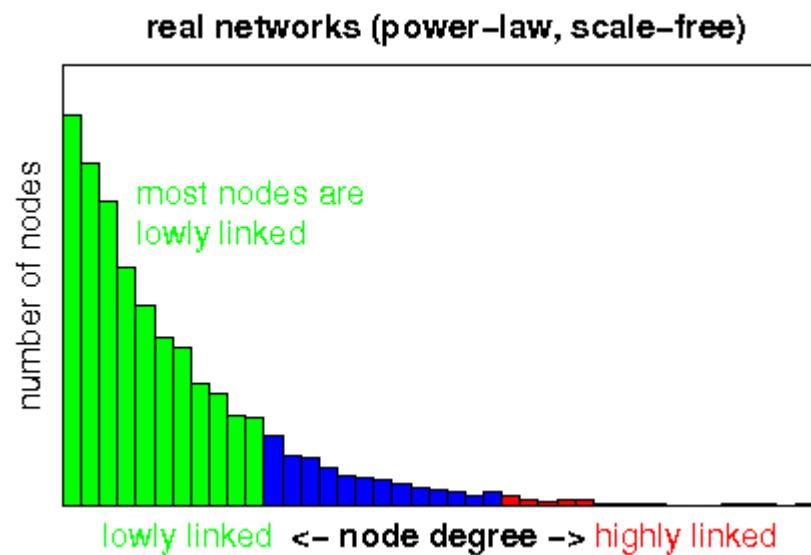
- Degree distribution follows a power law.

$$P(k) \sim k^{-\gamma}$$

$P(k)$ is the fraction of nodes that have k connections.

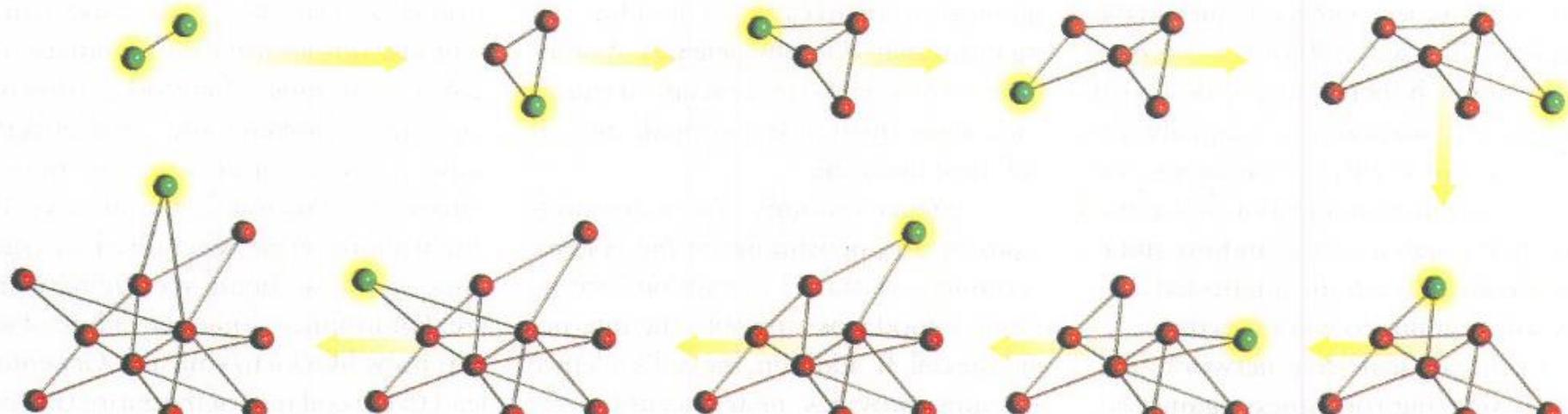
Typically, $2 < \gamma < 3$

- Number of nodes increases
- Preferential attachment
 - New nodes are preferentially attached to existing nodes with large numbers of connections
- Many real-world networks are thought to
 - World wide web
 - Social network
 - Citation network
- Barabási and Albert model

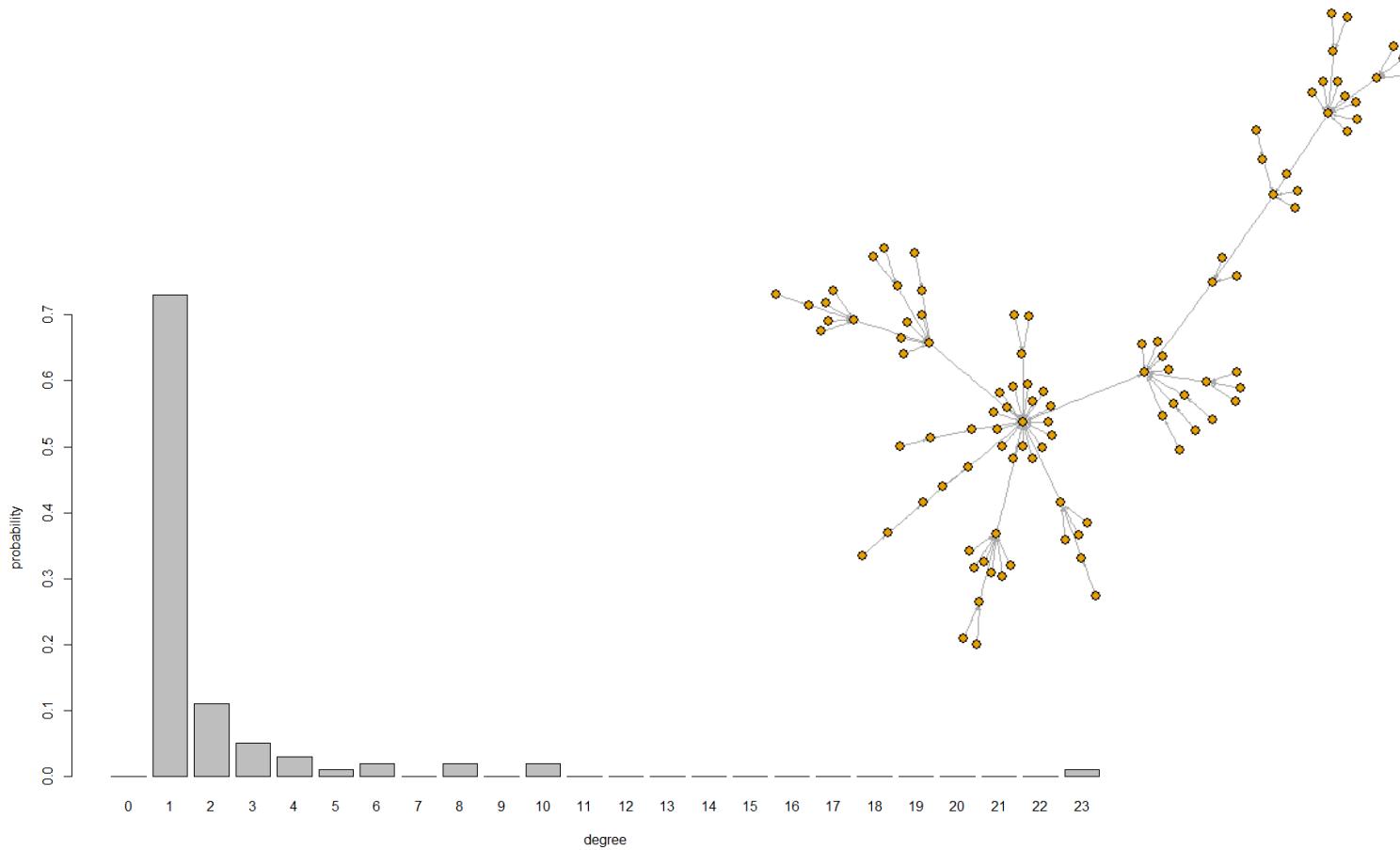


BIRTH OF A SCALE-FREE NETWORK

A SCALE-FREE NETWORK grows incrementally from two to 11 nodes in this example. When deciding where to establish a link, a new node (green) prefers to attach to an existing node (red) that already has many other connections. These two basic mechanisms—growth and preferential attachment—will eventually lead to the system's being dominated by hubs, nodes having an enormous number of links.



A Scale-Free Network



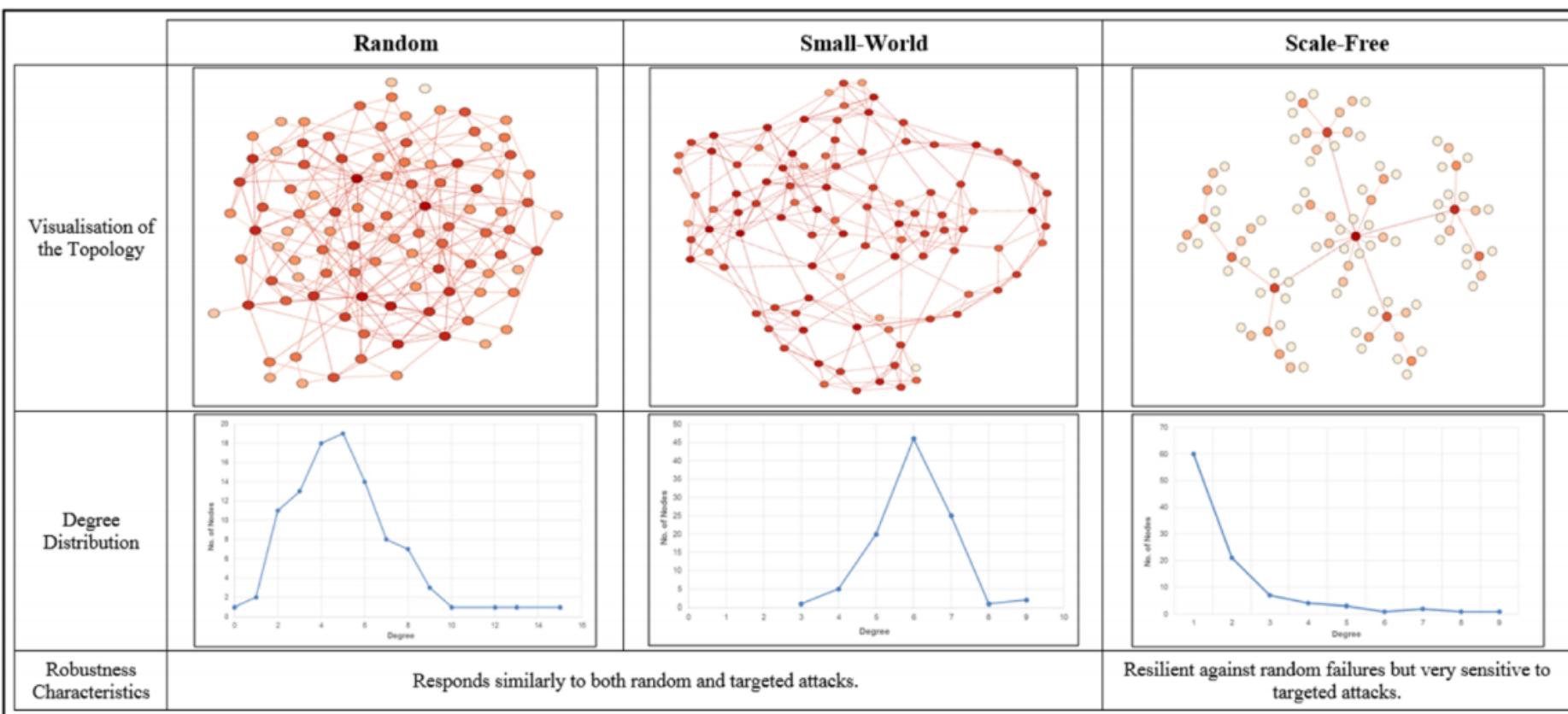


Fig. 1 Comparison of random, small-world and scale free networks. Topological structure of benchmark network models. Random and Small-world network topologies do not include hub nodes. In contrast, scale-free topologies are characterised by the presence of small number of highly connected hub nodes and a high number of feebly connected nodes. Presence of distinct hubs in scale-free networks make them more vulnerable to targeted attacks, compared to random and small-world networks

Centrality Measures

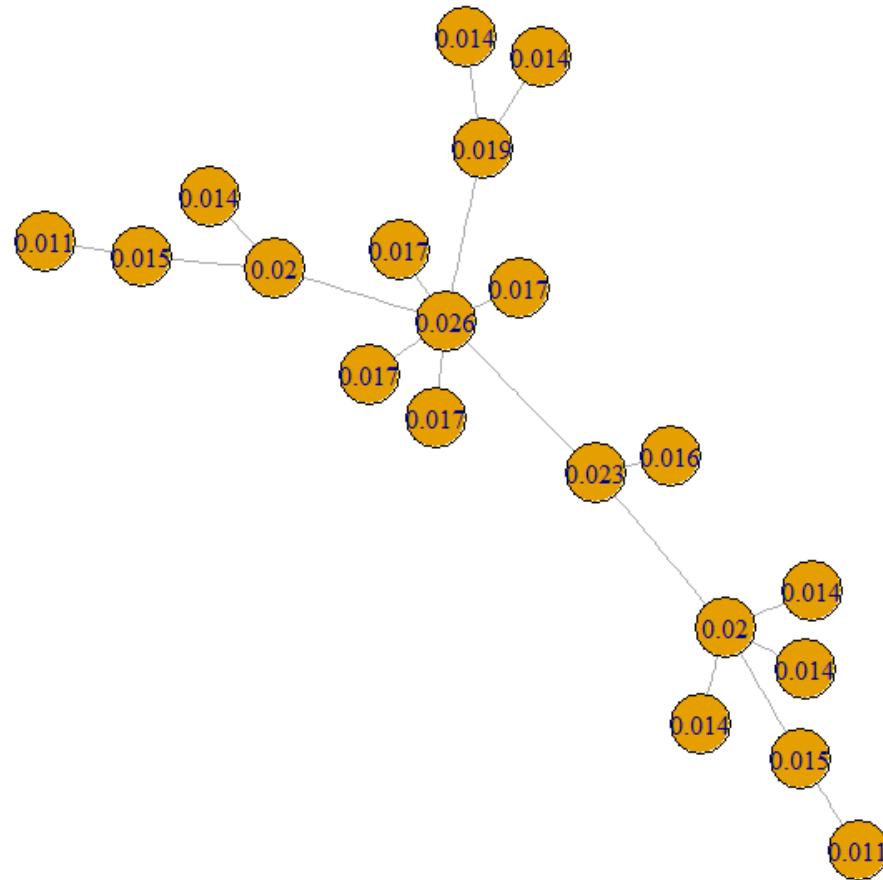
Network Centrality

- Which nodes in the network are most important?
Or most central?
- How centrality distributed among nodes?
↳ និមួយនឹង ឱ្យគ្រប់គ្រង
- Applications
 - Identify influencer in social network
 - If I am to pass on a message to three people in this network so that they in turn convey it to their friends and so on. Which three people should I select?
 - If I were to nominate a leader for a team of 500, whom should I pick?

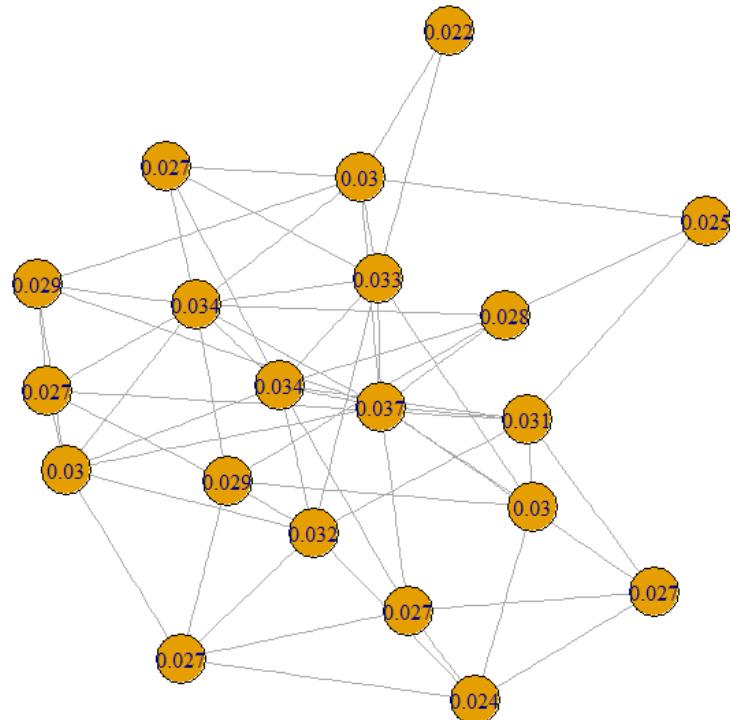
Centrality Measures

- Degree centrality
 - An important node is involved in large number of interactions.
- Closeness centrality *average path length*
 - An important node can communicate quickly with other nodes.
- Betweenness centrality
 - An important node will lie on most of the shortest paths between other nodes.
- Eigenvector centrality
 - An important node is connected to important neighbors.

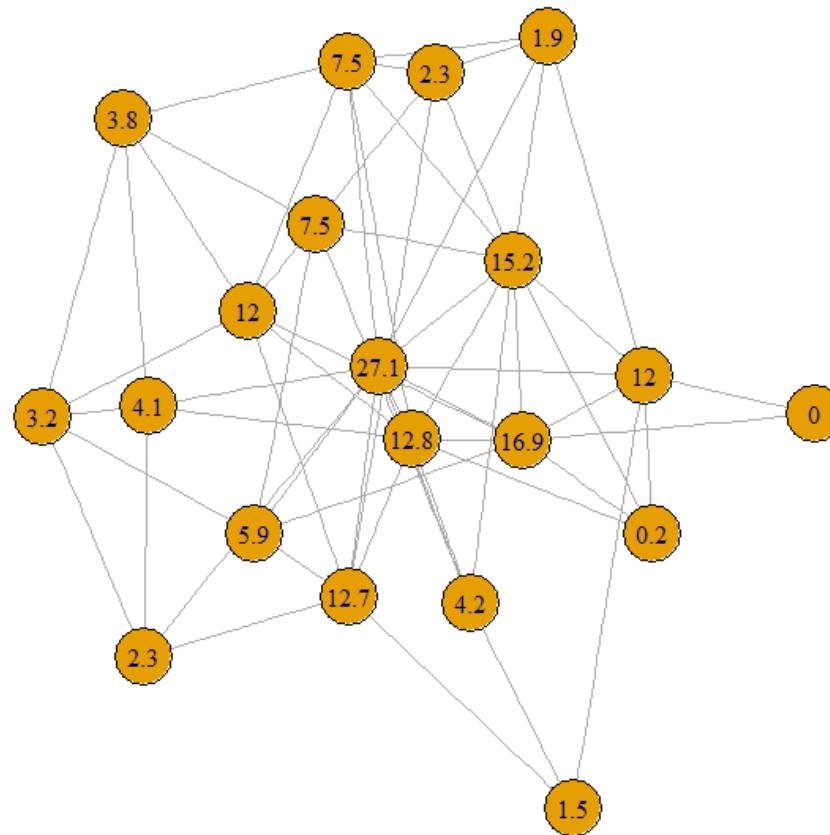
Closeness Centrality: scale-free network



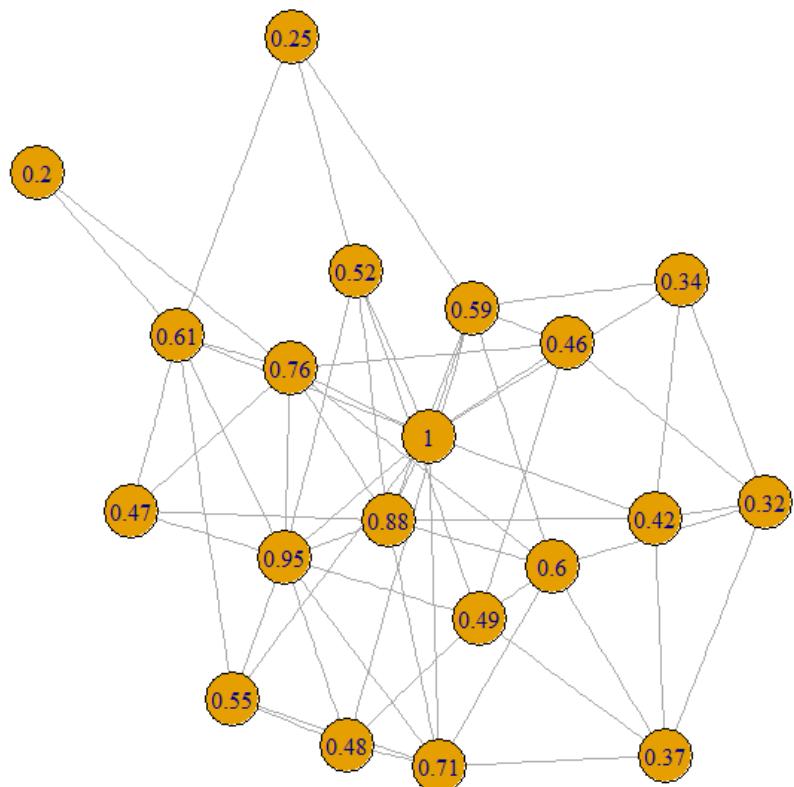
Closeness Centrality: small-world network



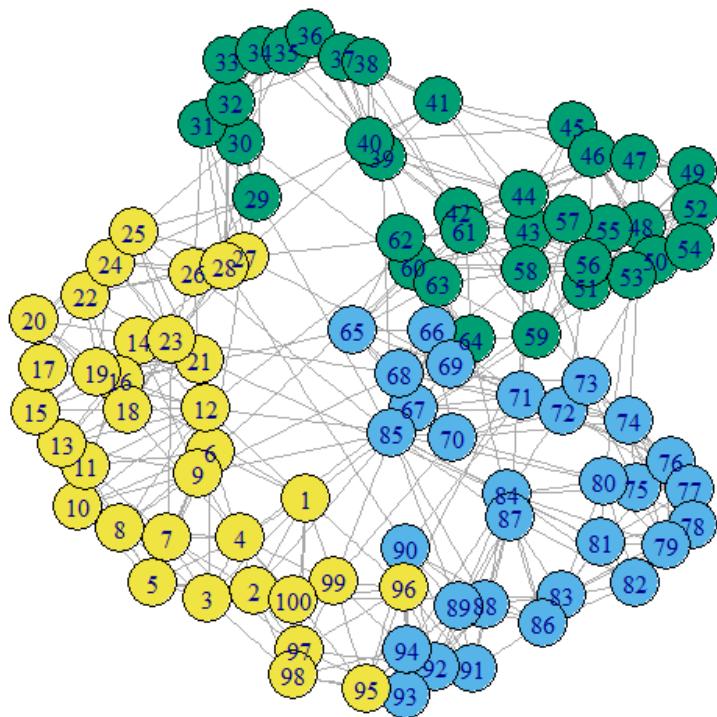
Betweenness Centrality: small-world network



Eigenvector Centrality : small-world network



Community Detection



Network Visualization

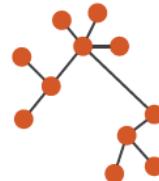
Network Visualization Styles

④ Node–Link Diagrams

Connection Marks

✓ NETWORKS

✓ TREES

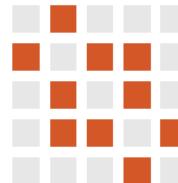


④ Adjacency Matrix

Derived Table

✓ NETWORKS

✓ TREES



④ Enclosure

Containment Marks

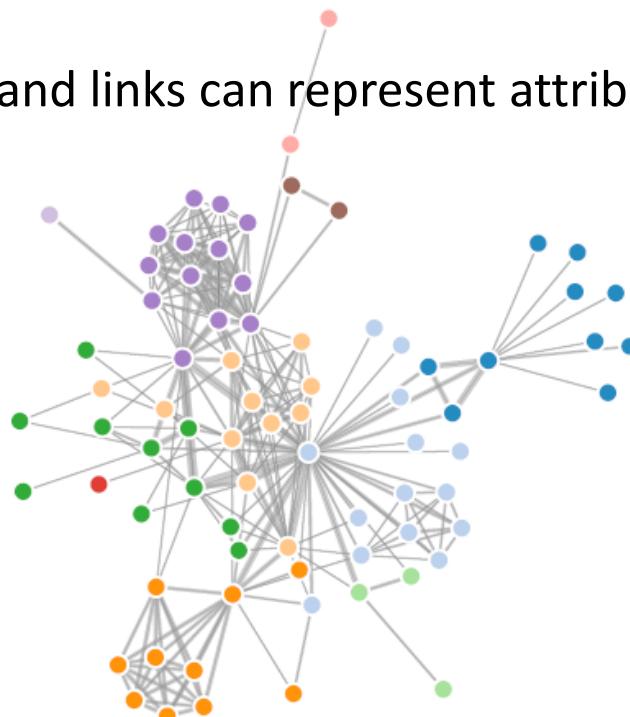
✗ NETWORKS

✓ TREES



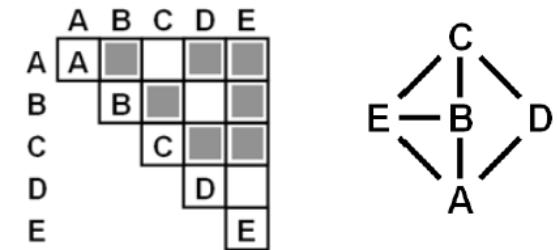
Node-Link Diagrams

- Link represents relation
- Good for exploring topology, locating clusters, following path
- Cluttering with high edge:node ratio
- Layout matters
- Size and color of nodes and links can represent attributes

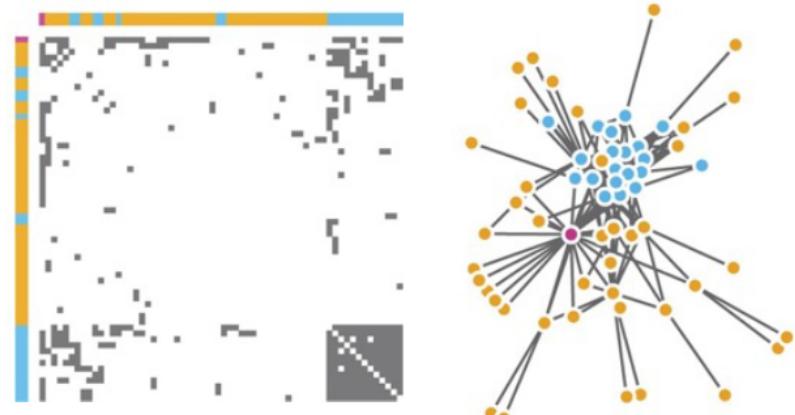
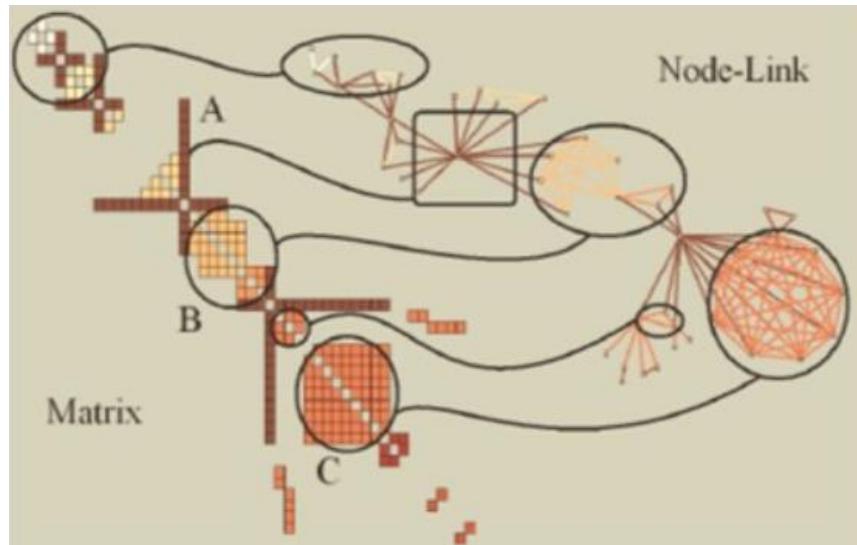


Adjacency Matrix

- Each marked cell represents an edge.
- Long straight lines represent high-degree nodes
- Hot spots represent clusters
- Good for dense network.
- Hard to follow paths.
- Node ordering matters.

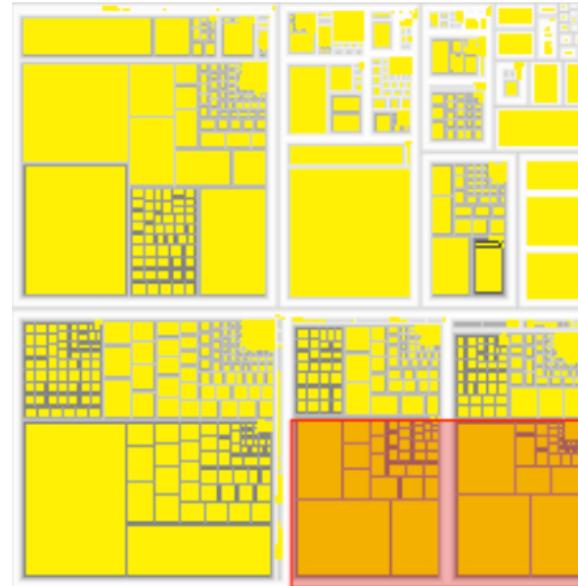


[NodeTrix: a Hybrid Visualization of Social Networks.
Henry, Fekete, and McGuffin. IEEE TVCG (Proc. InfoVis)
13(6):1302-1309, 2007.]



Enclosure

- TreeMap
- For tree only
- Sub-areas represent children nodes



Force-Directed Layout

- Spring model
 - Edges – spring (gravity, attraction)
 - Nodes – charged particles (repulsion)
- Iteratively recalculate positions of nodes according to forces on each node until the network is stable.

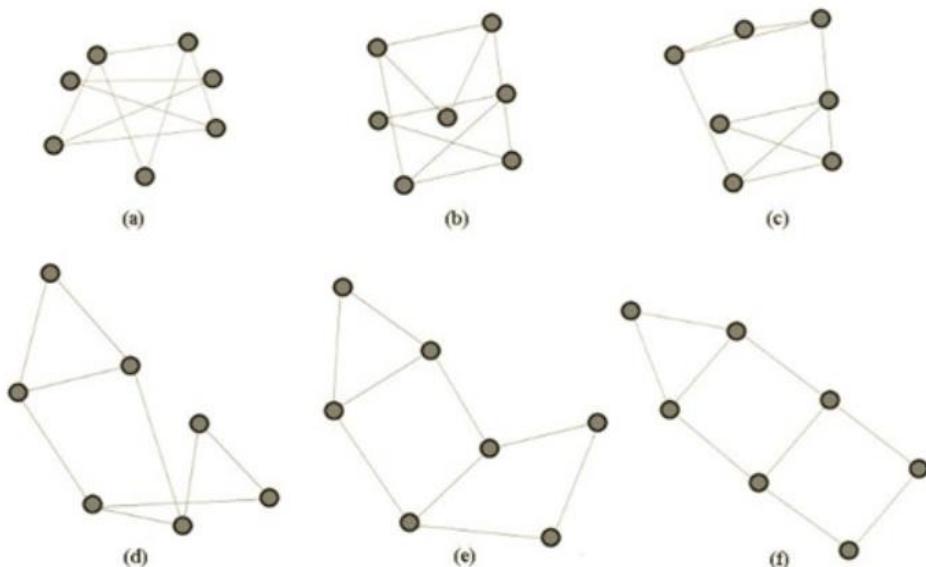
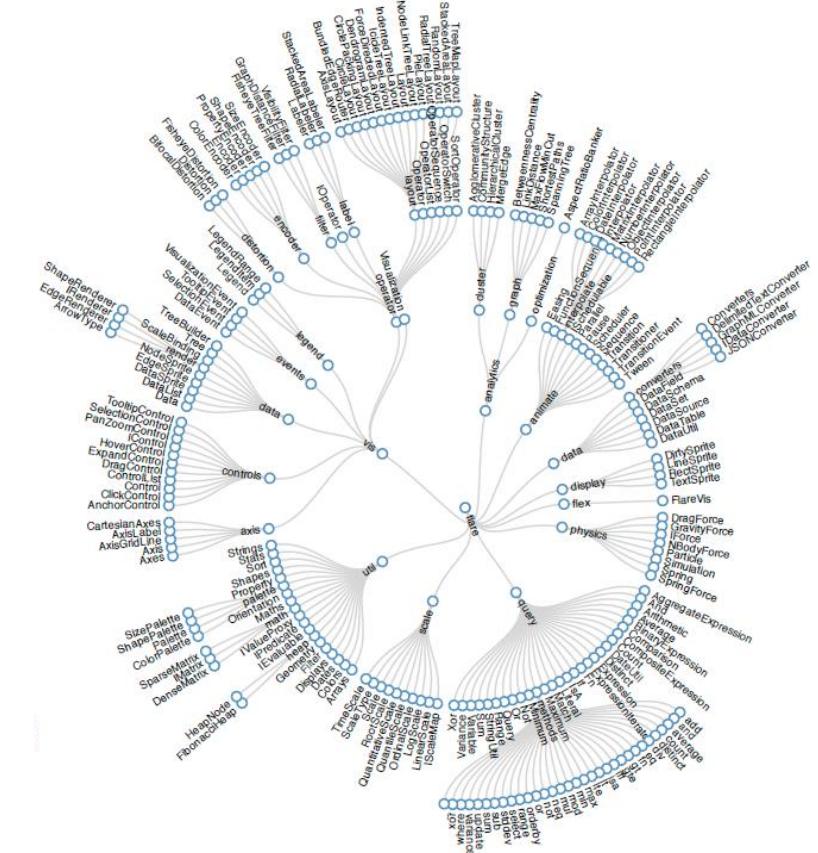
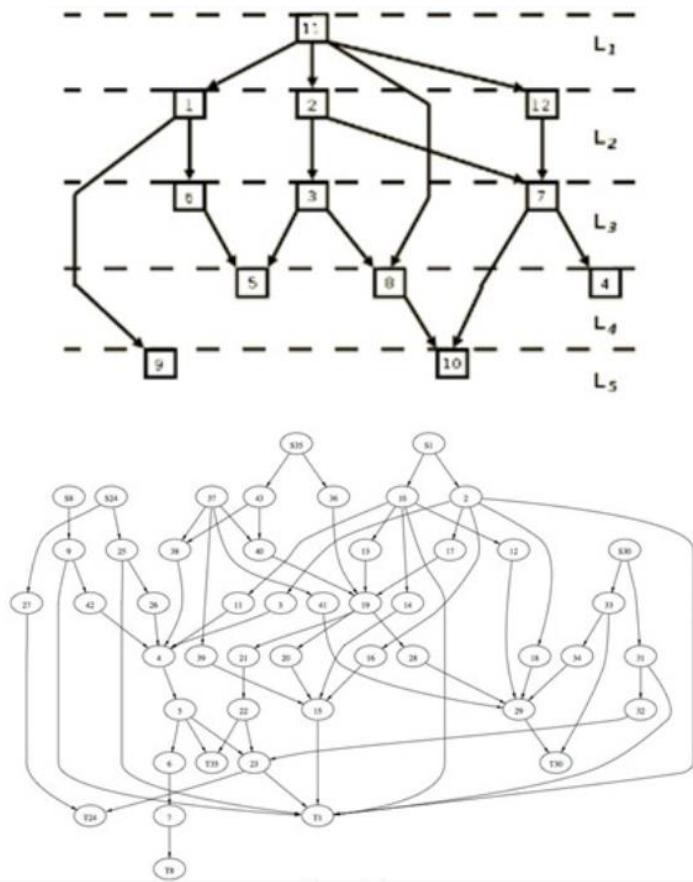


Figure 2: A graph drawing through a number of iterations of a force directed algorithm.

Hierarchical Layout

- Good for tree-like networks



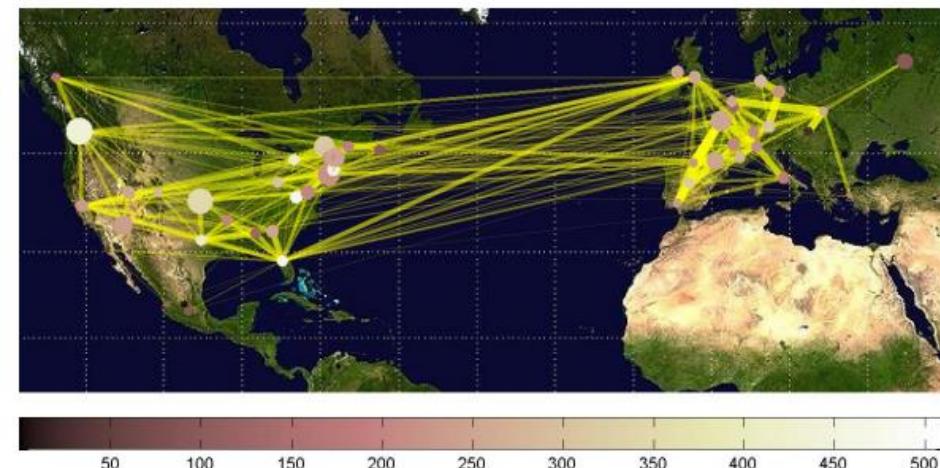
Circular Layout

- Nodes are positioned in a circle.
- Nodes with similar attributes may be grouped to see interactions between groups

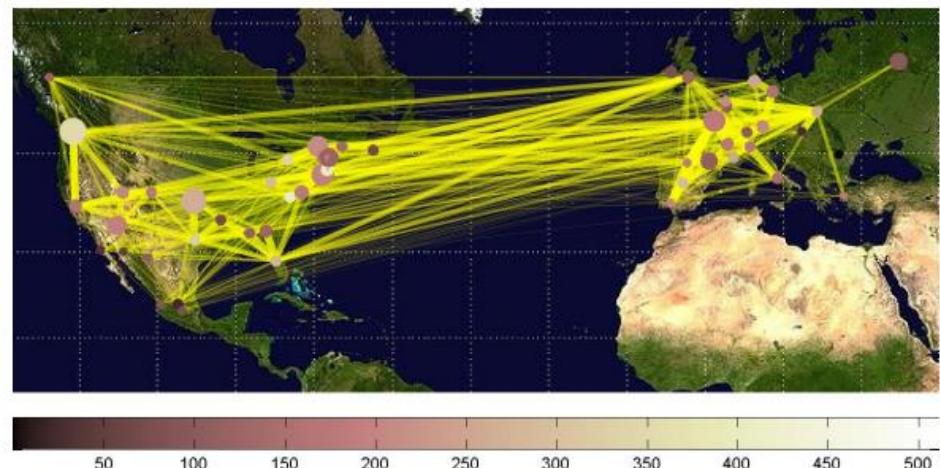


Network on Map

Figure 2a: Geographic map of buyer-supplier sub-network (the size of the bubble reflects location quotient, color gradient reflects the density of local linkages (normalized by the number of companies in the cluster) and the thickness of the line illustrates the density of trans-local linkages with other industrial clusters).



Buyer-supplier 2002-2005



Buyer-supplier 2010-2014

Network Analysis and Visualization with Gephi

Veera Muangsin

Network Analysis & Viz Tools

- Programs
 - Gephi
 - Cytoscape
 - <https://cytoscape.org/>
- Python Libraries
 - NetworkX
 - <https://networkx.org/>
 - igraph
 - <https://igraph.org/python/>
 - Plotly (visualization)
 - <https://plotly.com/python/network-graphs/>

Gephi

- Network Analysis and Visualization Tool
- https://gephi.org/tutorials/gephi-tutorial-quick_start.pdf
- <http://www.martingrandjean.ch/gephi-introduction/>

Sample datasets

- <https://github.com/gephi/gephi/wiki/Datasets>
- How to pronounce Gephi
 - Gef-fee
 - G-Fai



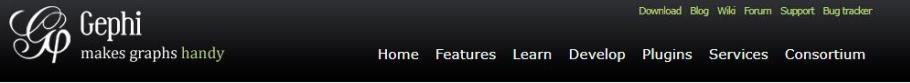
Gephi graph viz @Gephi · 4 Apr 2013

@jeanburgess @timhighfield @snurb_dot_info #Gephi should be pronounced /
gɛfi/ (how to read this: ow.ly/jKfqz) #TipOfTheDay



Download and Install Gephi

- <https://gephi.org>



The Open Graph Viz Platform

Gephi is the leading visualization and exploration software for all kinds of graphs and networks. Gephi is open-source and free.

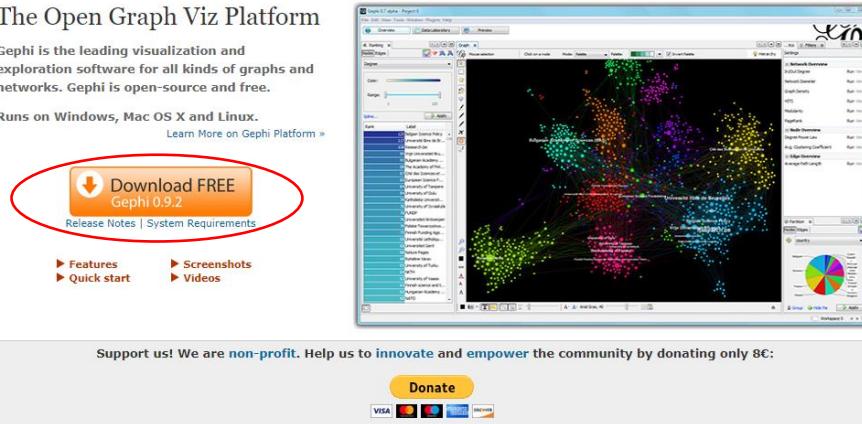
Runs on Windows, Mac OS X and Linux.

[Learn More on Gephi Platform >](#)

[Download FREE Gephi 0.9.2](#)

[Release Notes](#) | [System Requirements](#)

[Features](#) [Quick start](#) [Screenshots](#) [Videos](#)



Support us! We are non-profit. Help us to innovate and empower the community by donating only 8€:

[Donate](#)

VISA MASTERCARD PAYPAL

APPLICATIONS

- ✓ **Exploratory Data Analysis:** intuition-oriented analysis by networks manipulations in real time.
- ✓ **Link Analysis:** revealing the underlying structures of associations between objects.
- ✓ **Social Network Analysis:** easy creation of social data connectors to map community organizations and small-world networks.
- ✓ **Biological Network analysis:** representing patterns of biological data.
- ✓ **Poster creation:** scientific work promotion with hi-quality printable maps.

[Learn More >](#)

PAPERS

Like Photoshop™ for graphs. — the Community

LATEST NEWS

- Gephi updates with 0.9.2 version
- Gephi updates with 0.9.1 version
- A close look at the Gephi user community
- Gephi 0.9 released: Play with network data again
- Gephi boosts its performance with new "GraphStore" core

[See All >](#)

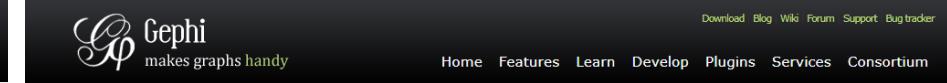
Tweets by @Gephi

 Gephi graph viz Retweeted

 Clement Levallais @seinecle

The @Gephi group on Facebook has passed

[Learn More >](#)



Download

Gephi is an open-source and multiplatform software distributed under the dual license [CDDL 1.0](#) and [GNU General Public License v3](#).

[Official Releases](#)

[Release Notes](#) | [System Requirements](#) | [Installation instructions](#)

Gephi 0.9.2 is the latest stable release.

[Download Gephi for Windows](#)

Version 0.9.2

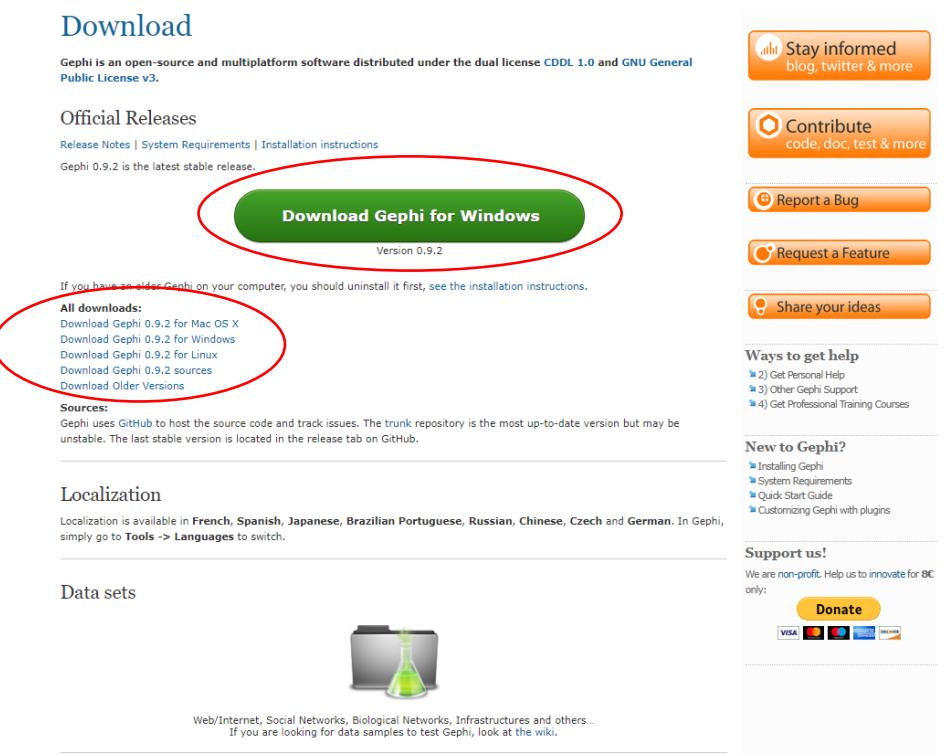
If you have an older Gephi on your computer, you should uninstall it first, see the installation instructions.

All downloads:

- [Download Gephi 0.9.2 for Mac OS X](#)
- [Download Gephi 0.9.2 for Windows](#)
- [Download Gephi 0.9.2 for Linux](#)
- [Download Gephi 0.9.2 sources](#)
- [Download Older Versions](#)

Sources:

Gephi uses GitHub to host the source code and track issues. The trunk repository is the most up-to-date version but may be unstable. The last stable version is located in the release tab on GitHub.



Localization

Localization is available in [French](#), [Spanish](#), [Japanese](#), [Brazilian Portuguese](#), [Russian](#), [Chinese](#), [Czech](#) and [German](#). In Gephi, simply go to [Tools -> Languages](#) to switch.

Data sets



Web/Internet, Social Networks, Biological Networks, Infrastructures and others...

If you are looking for data samples to test Gephi, look at the wiki.

Ways to get help

- 1) Get Personal Help
- 2) Other Gephi Support
- 3) Get Professional Training Courses

New to Gephi?

- [Installing Gephi](#)
- [System Requirements](#)
- [Quick Start Guide](#)
- [Customizing Gephi with plugins](#)

Support us!

We are non-profit. Help us to innovate for 8€ only:

[Donate](#)

VISA MASTERCARD PAYPAL

Learn how to use Gephi

See the Learn case that contains all the documentation to learn how to use Gephi, or jump directly to the Tutorial Quick Start!

Gephi requires Java

- Install Java
- If you run Gephi and find this error message

“Cannot find Java 1.8 or higher”

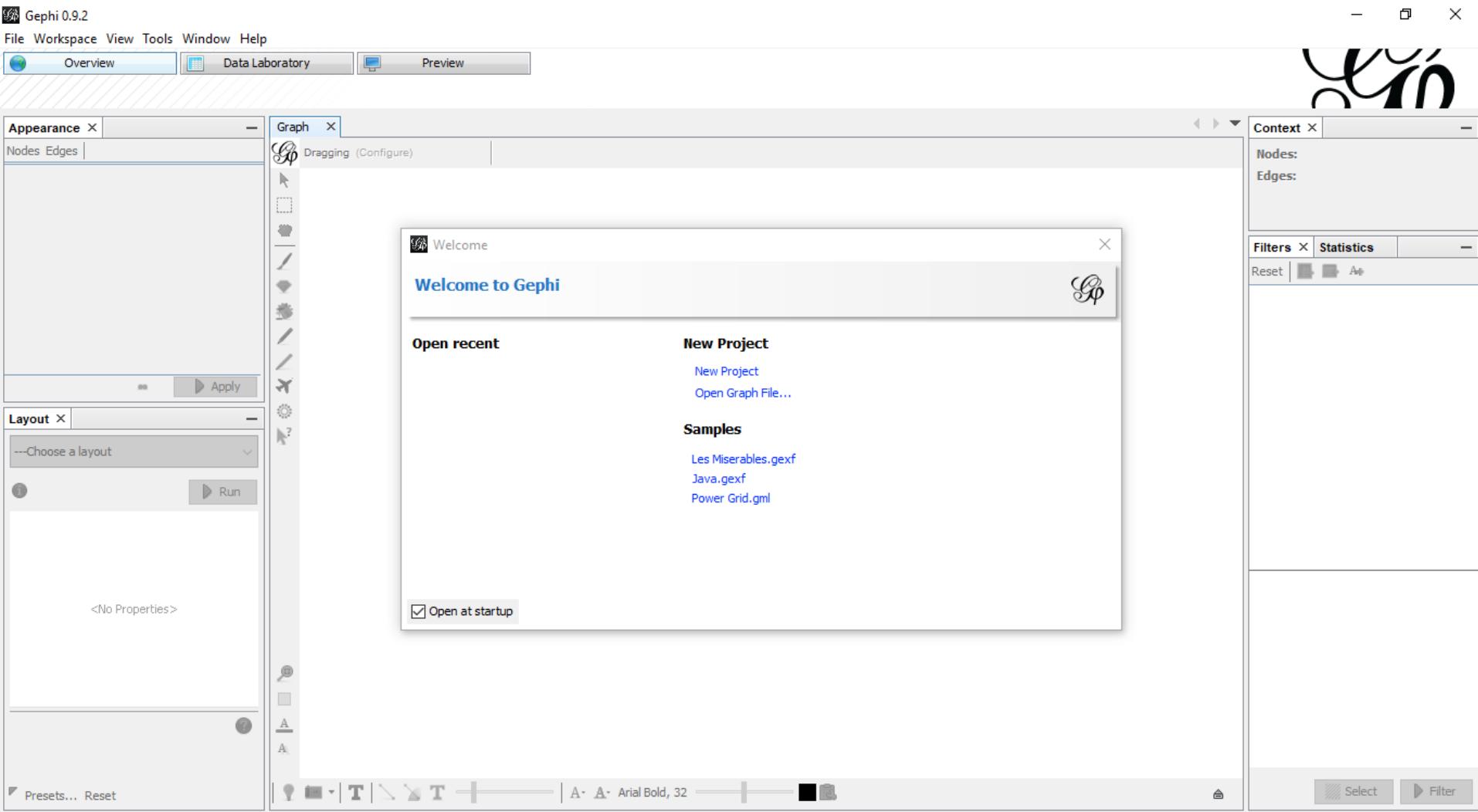
- Edit file `C:\Program Files\Gephi-0.9.2\etc\gephi.conf`
- Find line
`#jdkhome="/path/to/jdk"`
- Remove `#` and change `"/path/to/jdk"` to actual java directory.

For example:

`jdkhome="C:\Program Files (x86)\Java\jre1.8.0_211"`

Gephi First Page

Windows → Welcome





Sample: Les Misérables Network

"Do you hear the people sing?"

Co-occurrence Network: two characters are linked when they appear in the same chapter in the book.

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Appearance X

Nodes Edges

Graph X

Dragging (Configure)

Welcome

Welcome to Gephi

Open recent

netscience.gml

adjnoun.gml

New Project

New Project

Open Graph File...

Samples

Les Misérables.gexf

Java.gexf

Power Grid.gml

Open at startup

Layout X

--Choose a layout

Run

<No Properties>

Presets... Reset

Arial Bold, 32

Context X

Nodes:

Edges:

Filters Statistics X

Settings

Network Overview

Average Degree Run

Avg. Weighted Degree Run

Network Diameter Run

Graph Density Run

HITS Run

Modularity Run

PageRank Run

Connected Components Run

Node Overview

Avg. Clustering Coefficient Run

Eigenvector Centrality Run

Edge Overview

Avg. Path Length Run

Dynamic

Nodes Run

Edges Run

Degree Run

Clustering Coefficient Run

Network Size

Gephi 0.9.2

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Appearance X

Nodes Edges

Layout X

--Choose a layout Run

<No Properties>

Presets... Reset

Graph X

Dragging (Configure)

Import report

Source: Stream ImporterGEXF

Issues Report

Nodes	Issues
GEXF version 1.3	INFO

Graph Type: Undirected

of Nodes: 77

of Edges: 254

More options...

New workspace

Append to existing workspace

OK Cancel

Context X

Nodes:

Edges:

Filters Statistics X

Settings

Network Overview

- Average Degree Run
- Avg. Weighted Degree Run
- Network Diameter Run
- Graph Density Run
- HITS Run
- Modularity Run
- PageRank Run
- Connected Components Run

Node Overview

- Avg. Clustering Coefficient Run
- Eigenvector Centrality Run

Edge Overview

- Avg. Path Length Run

Dynamic

- # Nodes Run
- # Edges Run
- Degree Run

Arial Bold, 32

Zoom and Pan

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Appearance X

Nodes Edges

Unique Partition Ranking

#c0c0c0

Apply

Graph X

Dragging (Configure)

Layout X

--Choose a layout

Run

<No Properties>

Presets... Reset

Context X

Nodes: 77
Edges: 254
Undirected Graph

Filters Statistics X

Settings

Network Overview

Average Degree Run
Avg. Weighted Degree Run
Network Diameter Run
Graph Density Run
HITS Run
Modularity Run
PageRank Run
Connected Components Run

Node Overview

Avg. Clustering Coefficient Run
Eigenvector Centrality Run

Edge Overview

Avg. Path Length Run

Dynamic

Nodes Run
Edges Run
Degree Run

Scanning tool icon

• Scrolling wheel to zoom

• Press right button and pan

Center On graph:
Reset size and position

Arial Bold, 32

Node Labels

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Appearance X

Nodes Edges

Unique Partition Ranking

#c0c0c0

Apply

Layout X

--Choose a layout

Run

<No Properties>

Presets... Reset

Graph X Data Table X

Dragging (Configure)

Fameuil Blachevryfavourite Dahlia Listoli holomyes Zephine Fantine Perpetue Marguerite Simple Woman2 Woman1 Bamatabois Brixbridge Fauchelevent Cochevalier MotherInnocent Gribier MmeDeR MmeMabeuvre MmeBaptiste Myriel Count Geborand Countess DeLo Napoleon Bonapartier Champmerier

Boulatruelle Cladine Gueulemer Bab Thenardier Magnon Coseite Gillenorm Tounai Gillon LtGillenot Bar MlleVa

Label text settings

Nodes Edges Show properties

Select attributes to display as labels

Id Label Modularity Class

OK Cancel

Context X

Nodes: 77 Edges: 254 Undirected Graph

Filters X Statistics

Reset Library

- Attributes
- Dynamic
- Edges
- Operator
- Topology
- Saved queries

Queries Drag filter here

Presets... Reset

Select Filter

toggle label

set label

Toggle Edges

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Appearance X

Nodes Edges

Unique Partition Ranking

#c0c0c0

Layout X

--Choose a layout

Run

<No Properties>

Presets... Reset

Graph X

Mouse selection (Configure)

Context X

Nodes: 77
Edges: 254
Undirected Graph

Filters X Statistics

Reset Library

- Attributes
- Dynamic
- Edges
- Operator
- Topology
- Saved queries

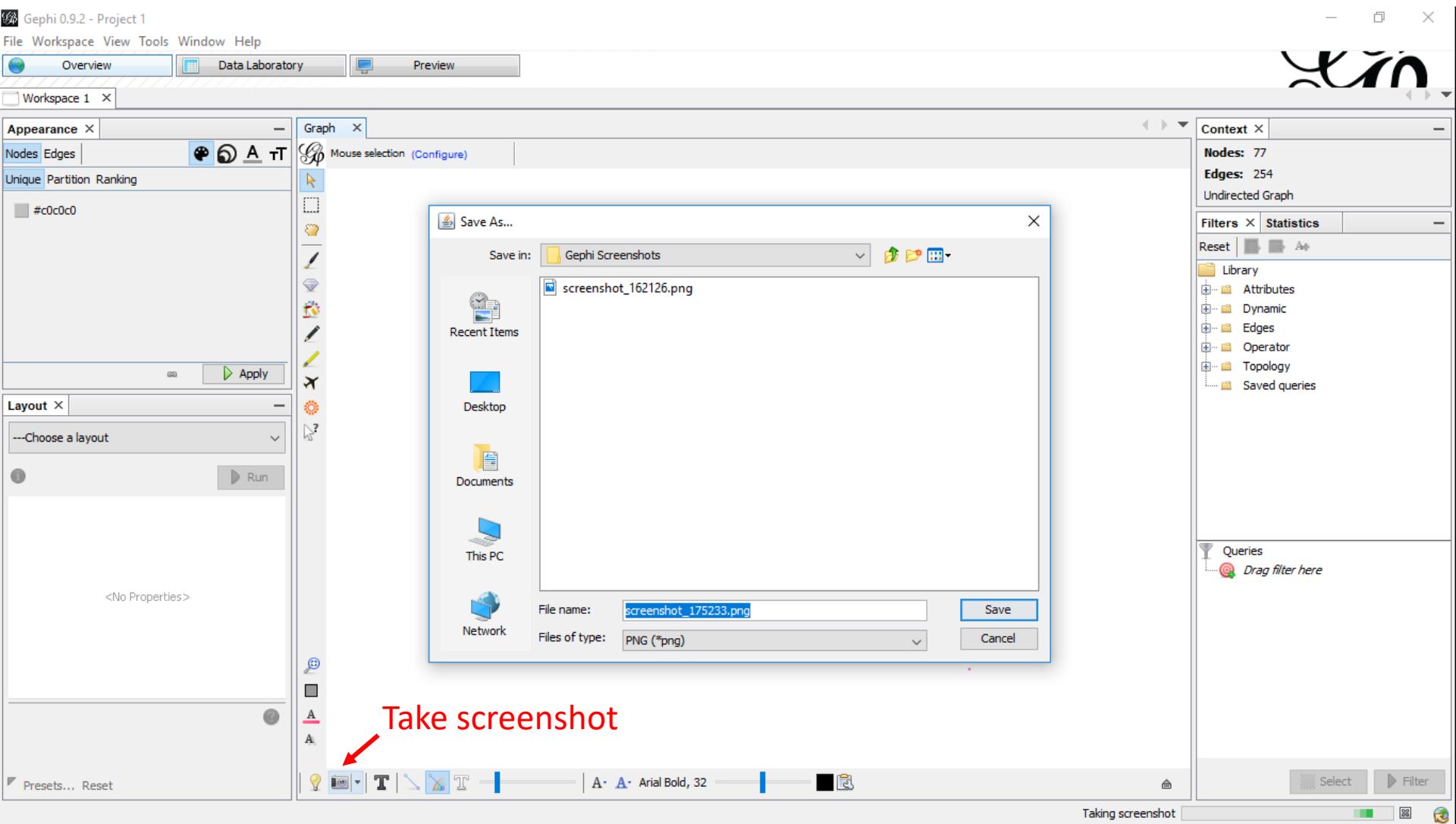
Queries

Drag filter here

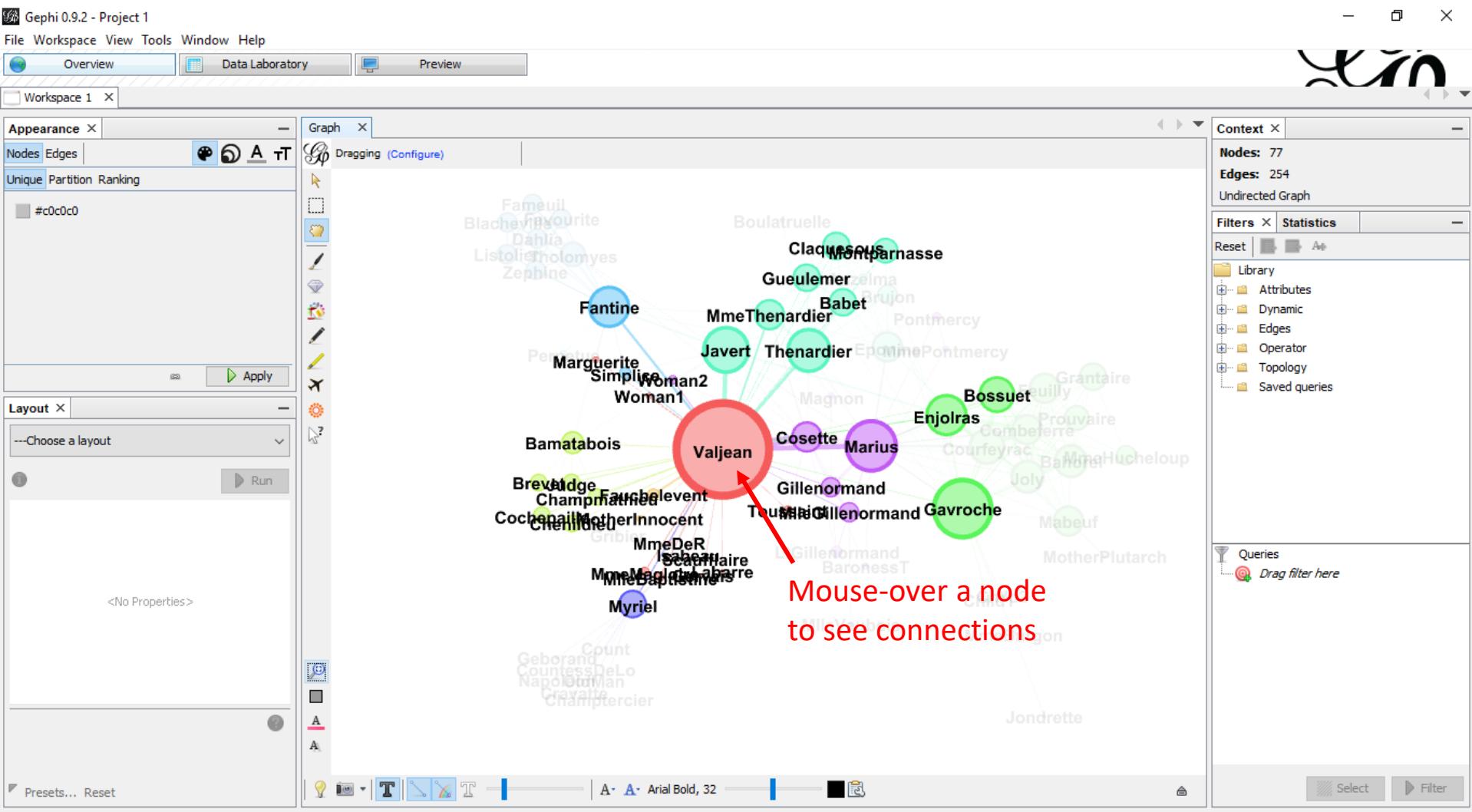
toggle edge

The screenshot displays the Gephi 0.9.2 software interface. The main workspace shows a network graph with numerous nodes of varying sizes and colors (red, blue, green, purple) scattered across the screen. A prominent large red node is centered. The left side of the interface features the 'Appearance' panel with 'Nodes Edges' selected, and the 'Layout' panel showing a dropdown menu for choosing a layout. The right side includes the 'Context' panel displaying statistics (Nodes: 77, Edges: 254, Undirected Graph), the 'Filters' panel, and the 'Library' panel which lists categories like Attributes, Dynamic, Edges, Operator, Topology, and Saved queries. A red arrow points to the 'Edges' icon in the bottom toolbar, with the text 'toggle edge' written above it.

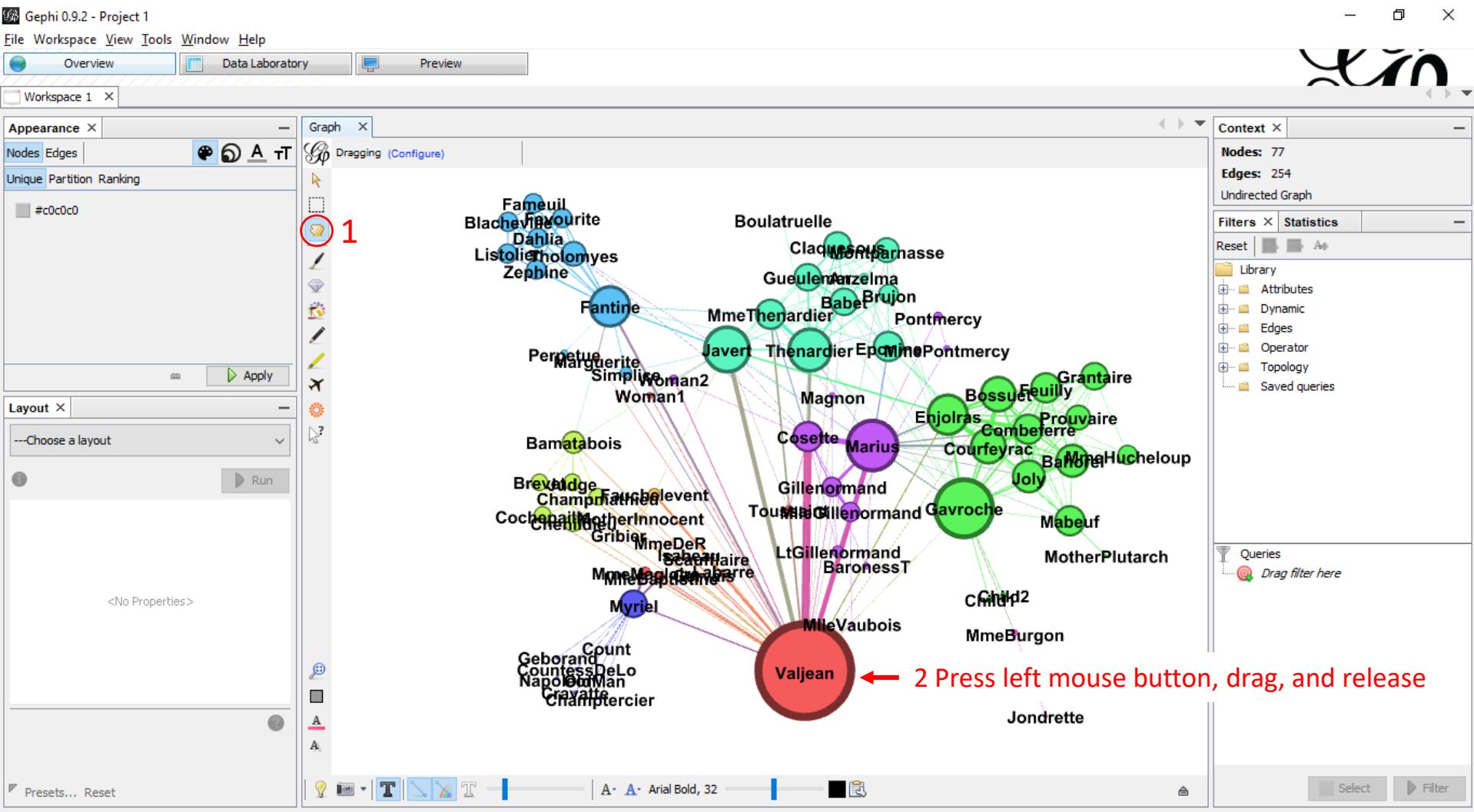
Take Screenshot



See connections of a node

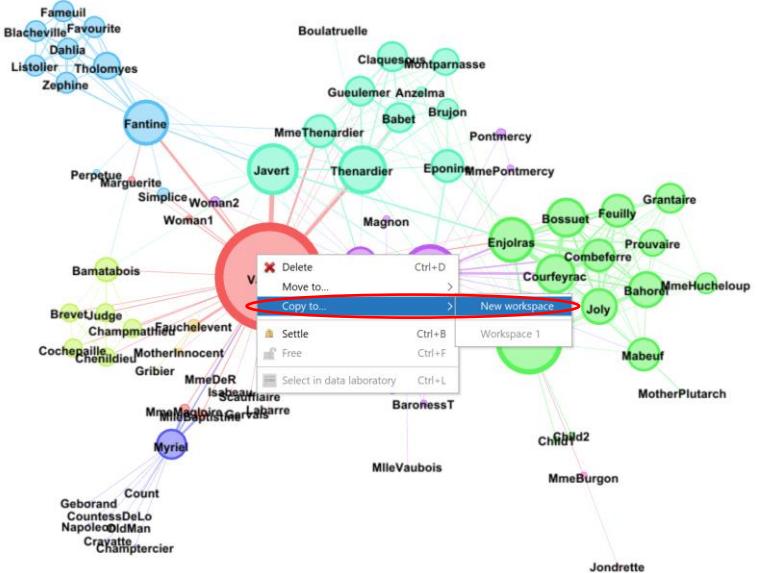
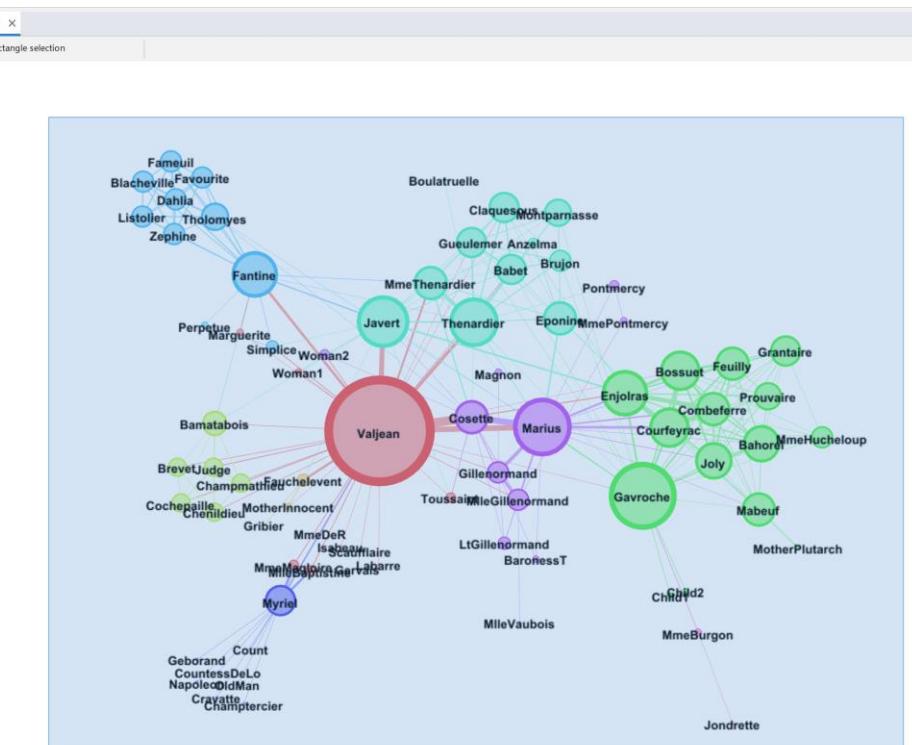


Move a node



Copy network to a new workspace

- Gephi doesn't have "undo/redo" feature. 😞
- To work around, before making changes, copy the network to a new workspace.
- Select nodes → right click → "Copy to..." → "New workspace"
- An alternative is to save versions.



Delete a node

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Appearance X

Nodes Edges Unique Partition Ranking

#c0c0c0

Apply

Layout X

--Choose a layout Run

<No Properties>

Presets... Reset

Graph X

Dragging (Configure)

Context X

Nodes: 77
Edges: 254
Undirected Graph

Filters X Statistics

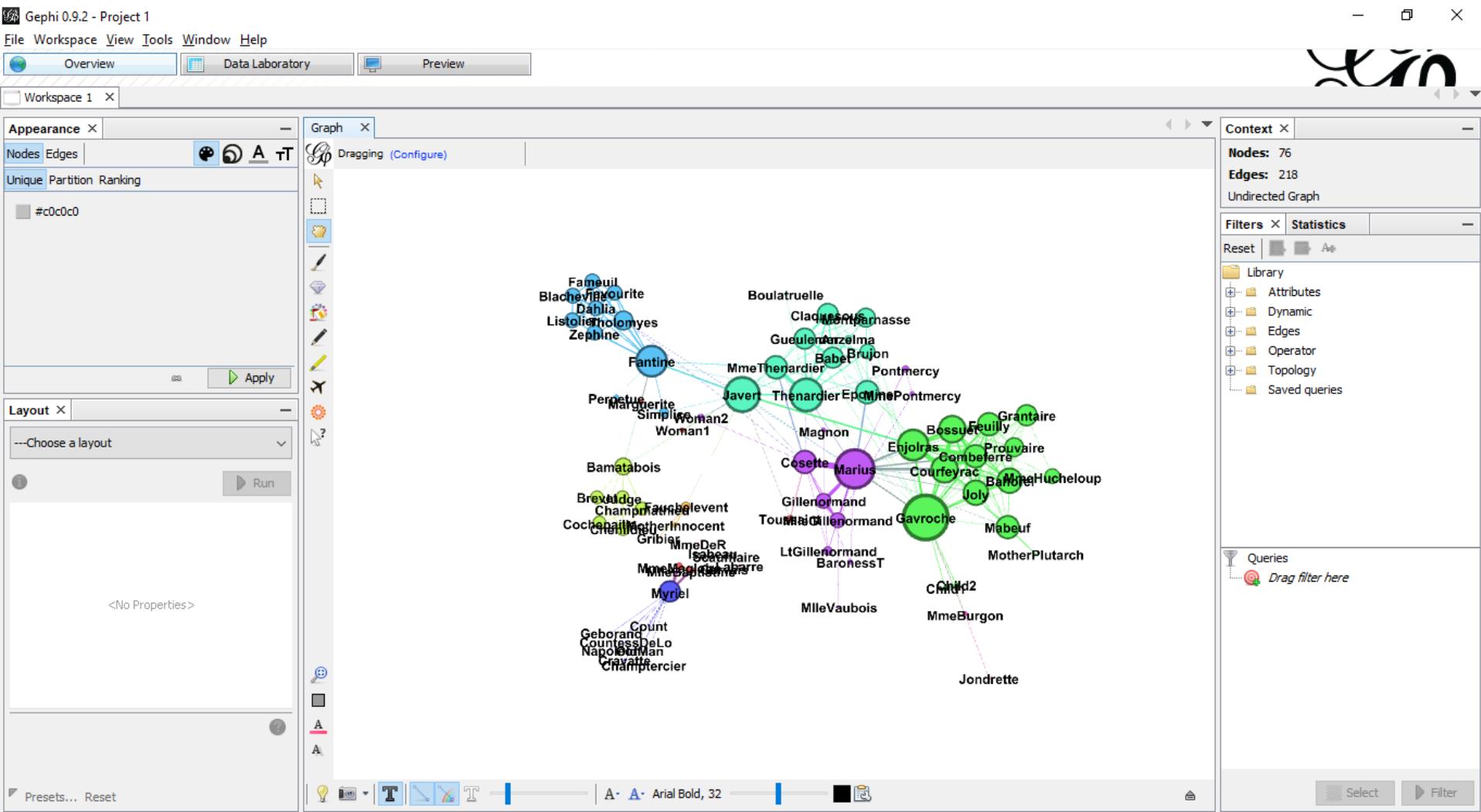
Reset Library Attributes Dynamic Edges Operator Topology Saved queries

Queries Drag filter here

Right click and delete

The screenshot shows a network graph in Gephi representing characters from Victor Hugo's Les Misérables. The nodes are colored according to their community structure. The central node, 'Valjean', is highlighted with a large red circle. A context menu is open over this node, listing options such as 'Delete', 'Move to...', 'Copy to...', 'Settle', 'Free', and 'Select in data laboratory'. A red arrow points specifically to the 'Delete' option. The background shows other nodes like 'Fantine', 'Cosette', 'Marius', and 'Javert' clustered together. The interface includes various panels for appearance, layout, and context management.

Delete a node: result



Resizable select tool

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Appearance X

Nodes Edges Unique Partition Ranking

#c0c0c0

Layout X

--Choose a layout Run

<No Properties>

Presets... Reset

Graph X

Mouse selection (Configure)

Context X

Nodes: 77
Edges: 254
Undirected Graph

Filters X Statistics

Reset Library Attributes Dynamic Edges Operator Topology Saved queries

Queries Drag filter here

Select Filter

Ctrl + mouse wheel

Edit a node

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Appearance Edit X

Valjean - Properties

- Size: 100.0
- Position (x): -74.03141
- Position (y): 12.371608
- Position (z): 0.0
- Color: [245,91,91]
- Label Size: 1.0
- Label Color: null
- Label Visible:

Valjean - Attributes

- Id: 11
- Label: Valjean

Layout X

--Choose a layout

Run

1

Graph X

Mouse selection (Configure)

3 modify data

2

Context X

Nodes: 77
Edges: 254
Undirected Graph

Filters X Statistics

Reset

Library

- Attributes
- Dynamic
- Edges
- Operator
- Topology
- Saved queries

Queries

Drag filter here

Select Filter

Presets... Reset

Arial Bold, 32

Data Laboratory

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Data Table

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

11	Id	Label
48		Valjean
55		Gavroche
27		Marius
25		Javert
23		Thenardier
58		Fantine
62		Enjolras
64		Courfeyrac
63		Bossuet
65		Bahorel
24		Joly
26		MmeThenardier
41		Cosette
57		Eponine
59		Mabeuf
61		Combeferre
0		Feuilly
66		Myriel
68		Grantaire
69		Gueulemer
70		Babet
16		Claquesous
60		Tholomyes
71		Prouvaire
29		Montparnasse
		Ramatabaric

Add column Merge columns Delete column Clear column Copy data to other column Fill column with a value

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Data Table

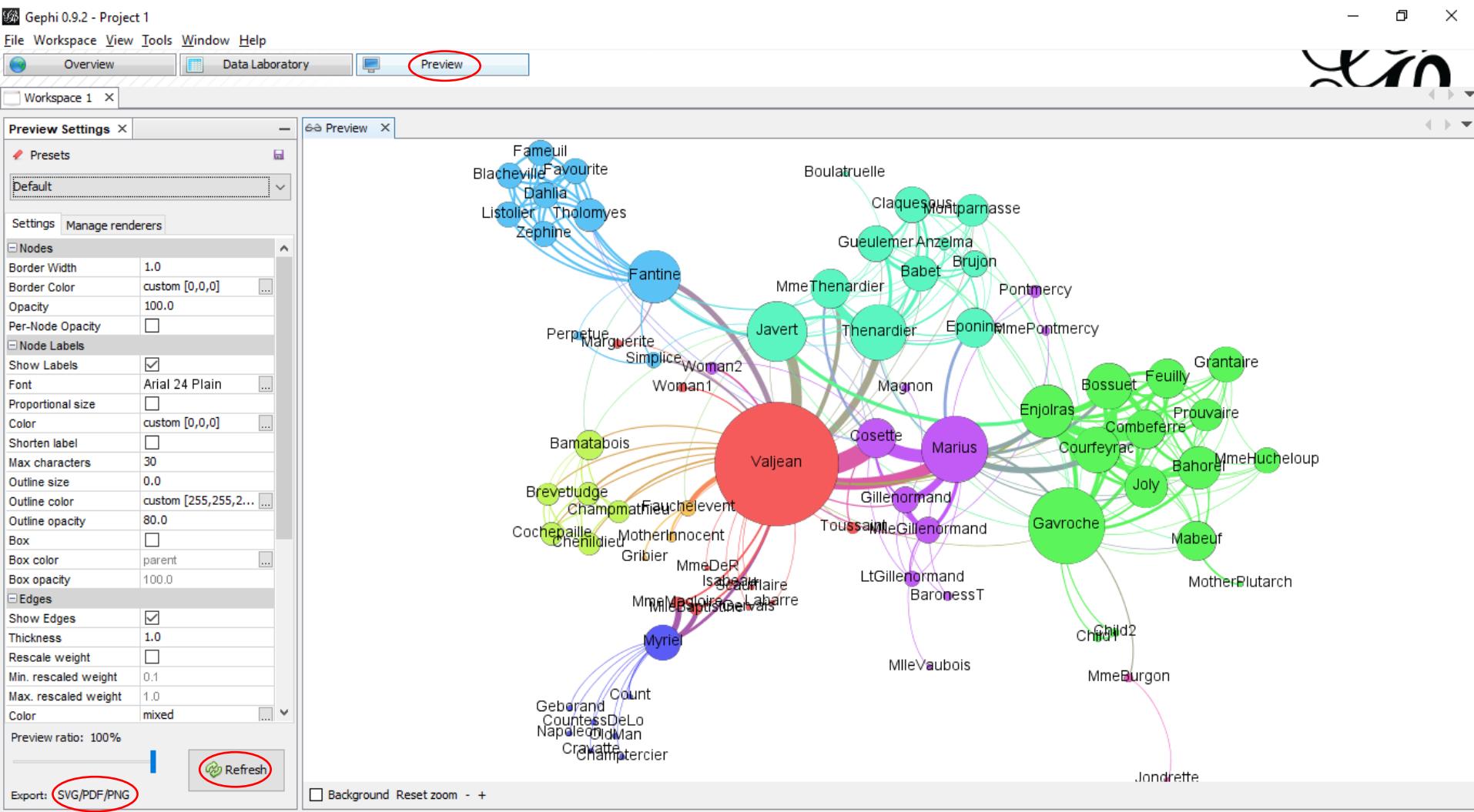
Nodes Edges Configuration Add node Add edge Search/Replace Import Spreadsheet Export table More

	Source	Target	Type	Id
1	ID of source node	0	ID of target node	Undirected 0
2		0		Undirected 1
3		0		Undirected 2
3		2		Undirected 3
4		0		Undirected 4
5		0		Undirected 5
6		0		Undirected 6
7		n		Undirected 7
				Undirected 8
				Undirected 9
				Undirected 13
11		2		Undirected 12
11		3		Undirected 11
11		10		Undirected 10
12		11		Undirected 14
13		11		Undirected 15
14		11		Undirected 16
15		11		Undirected 17
17		16		Undirected 18
18		16		Undirected 19
18		17		Undirected 20
19		16		Undirected 21
19		17		Undirected 22
19		18		Undirected 23
20		16		Undirected 24
20		17		Undirected 25

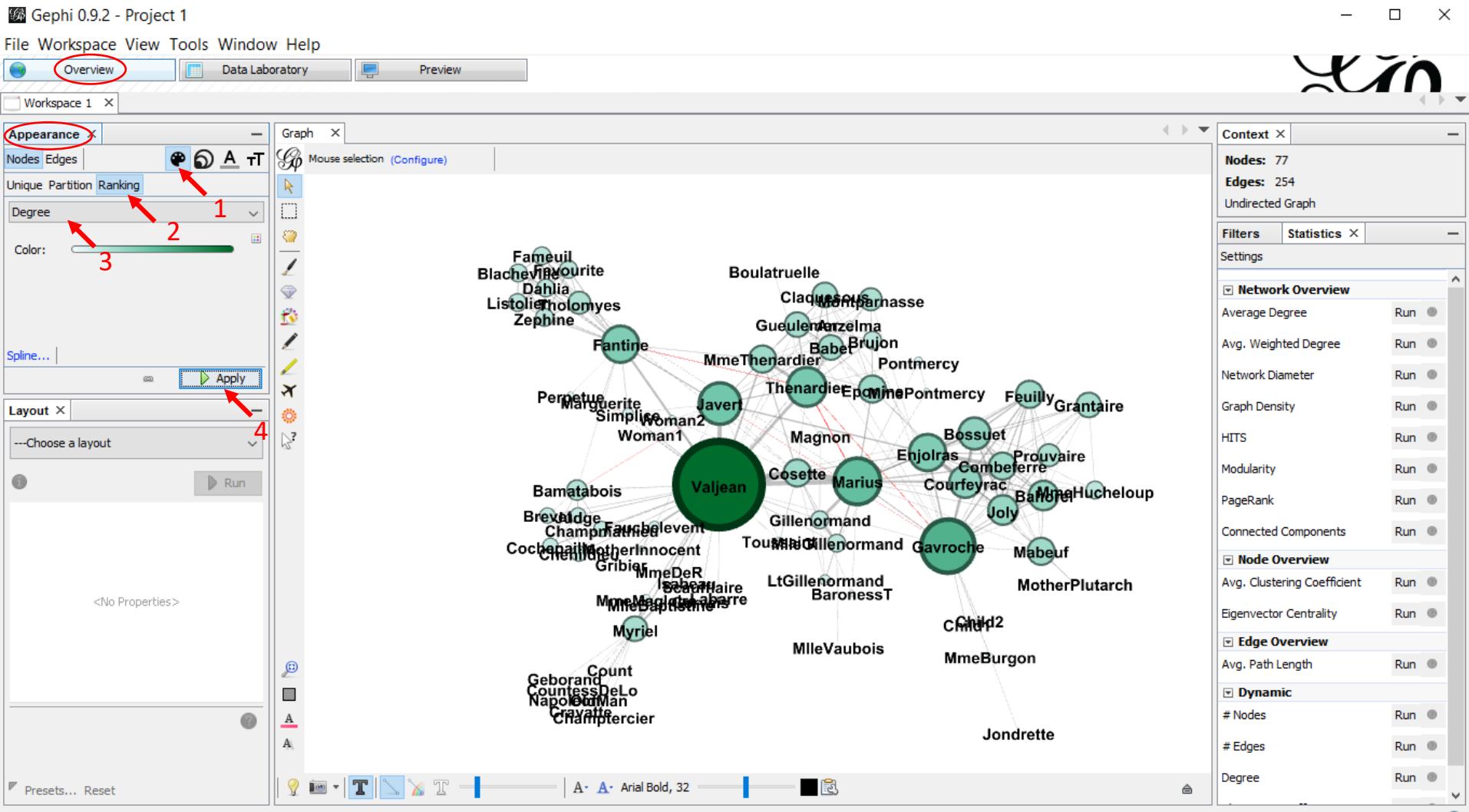
Add column Merge columns Delete column Clear column Copy data to other column Fill column with a value Duplicate column

Values in Data Table can be modified.

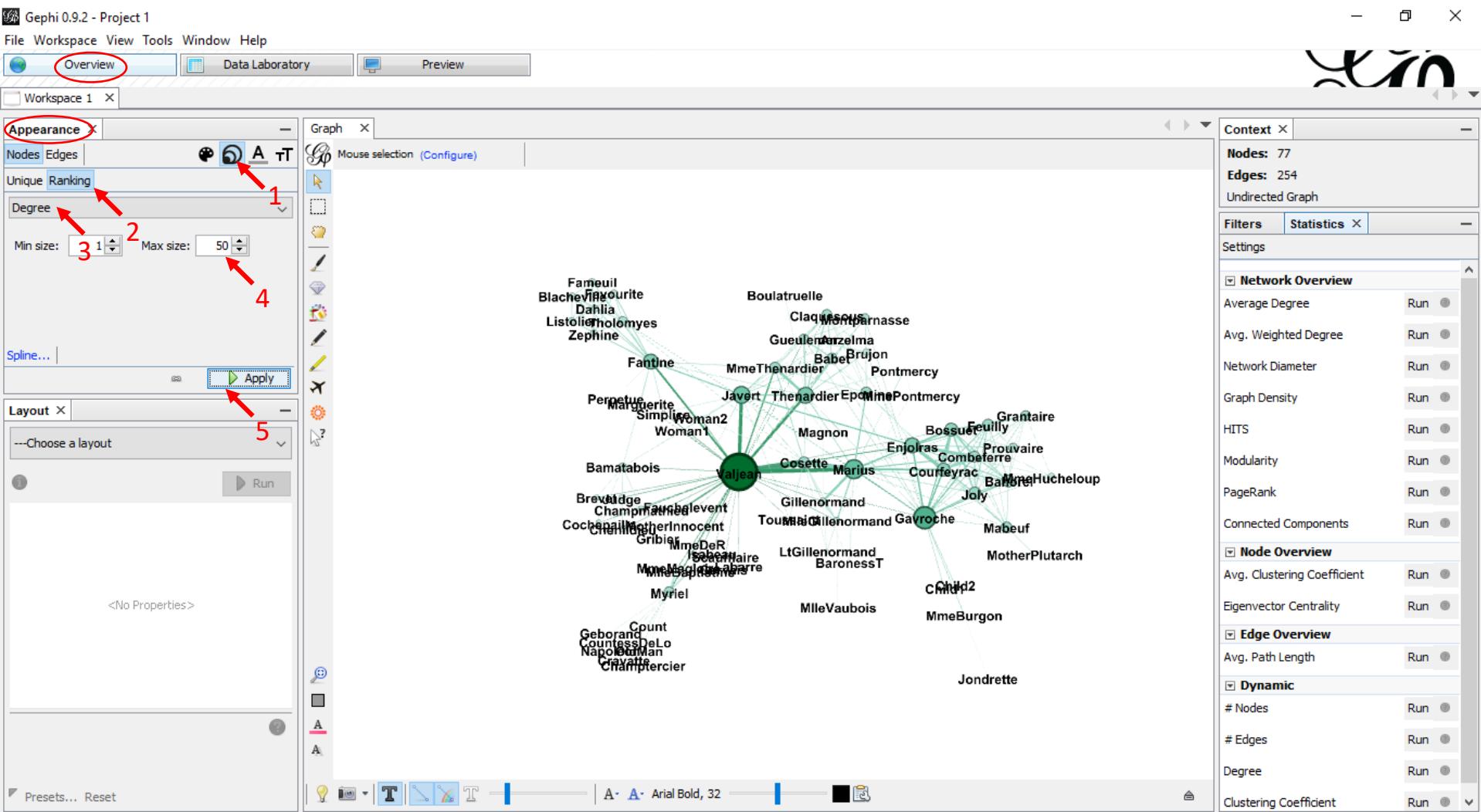
Preview



Color Nodes by Degrees



Rank Node Size by Degree



Statistics: Degree

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Appearance X

Nodes Edges Unique Ranking

Degree

Min size: 1 Max size: 49.5

Spline... Apply

Layout X

--Choose a layout Run

<No Properties>

Presets... Reset

Graph X

Mouse selection (Cont)

HTML Report

Degree Report

Results:

Average Degree: 6.597

Degree Distribution

Value	Count
2	17
4	6
6	5
8	1
10	5
12	2
14	1
16	1
18	1
22	1
36	1

Print Copy Save Close

Arial Bold, 32

Context X

Nodes: 77 Edges: 254 Undirected Graph

Filters Statistics

Network Overview

Average Degree 6.597 Run

Avg. Weighted Degree Run

Network Diameter Run

Graph Density Run

HITS Run

Modularity Run

PageRank Run

Connected Components Run

Node Overview

Avg. Clustering Coefficient Run

Eigenvector Centrality Run

Edge Overview

Avg. Path Length Run

Dynamic

Nodes Run

Edges Run

Degree Run

Print Copy Save Close

Arial Bold, 32

Presets... Reset

Graph Distance and Centrality Measures

The screenshot shows the Gephi 0.9.2 interface with the 'Graph Distance settings' dialog box open. The dialog box contains information about graph distance, distance types (Directed vs. Undirected), and centrality measures (Betweenness, Closeness, Eccentricity). The 'OK' button is highlighted with a red arrow labeled '1'. The background shows the Gephi workspace with various tools and panels like Appearance, Layout, and Context.

Graph Distance settings

Distance
The average graph-distance between all pairs of nodes. Connected nodes have graph distance 1. The diameter is the longest graph distance between any two nodes in the network. (i.e. How far apart are the two most distant nodes).

Directed Normalize Centralities in [0,1]
 Undirected

Betweenness Centrality: Measures how often a node appears on shortest paths between nodes in the network.
Closeness Centrality: The average distance from a given starting node to all other nodes in the network.
Eccentricity: The distance from a given starting node to the farthest node from it in the network.

OK Cancel

Note: This can take a long time for a large network.

1 2

Gephi 0.9.2 - Project 1
File Workspace View Tools Window Help
Overview Data Laboratory Preview
Workspace 1
Appearance X
Nodes Edges Unique Ranking
Degree Min size: 1 Max size: 50
Spline... Apply
Layout X
---Choose a layout Run
<No Properties>
Presets... Reset
Graph X
Mouse selection (Configure)
Context X
Nodes: 77 Edges: 254 Undirected Graph
Filters Statistics X
Settings
Network Overview
Average Degree 6.597 Run
Avg. Weighted Degree 21.299 Run
Network Diameter Run 1
Graph Density Run
HITS Run
Modularity Run
PageRank Run
Connected Components Run
Node Overview Avg. Clustering Coefficient Run
Eigenvector Centrality Run
Edge Overview Avg. Path Length Run
Dynamic # Nodes Run
Edges Run
Degree Run
Clustering Coefficient Run

Modularity

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory

Workspace 1

Appearance X

Nodes Edges Unique Ranking

Degree

Min size: 1 Max size: 50

Spline... Apply

Layout X

--Choose a layout Run

<No Properties>

Presets... Reset

Graph X

Mouse

Appearance X

Nodes: 77 Edges: 254 Undirected Graph

Filters Statistics

Context X

Nodes: 77 Edges: 254 Undirected Graph

Network Overview

Average Degree 6.597 Run

Avg. Weighted Degree 21.299 Run

Network Diameter 5 Run

Graph Density Run

HITS Run

Modularity 0.567 Run

PageRank Run

Connected Components Run

Node Overview

Avg. Clustering Coefficient Run

Eigenvector Centrality Run

Edge Overview

Avg. Path Length 2.641 Run

Dynamic

Nodes Run

Edges Run

Degree Run

Clustering Coefficient Run

HTML Report

Modularity Report

Parameters:

Randomize: On
Use edge weights: On
Resolution: 1.0

Results:

Modularity: 0.567
Modularity with resolution: 0.567
Number of Communities: 6

Size Distribution

Modularity Class	Size (number of nodes)
0	11
1	17
2	11
3	22
4	10
5	6

Print Copy Save Close

Statistic results

New columns are added.

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

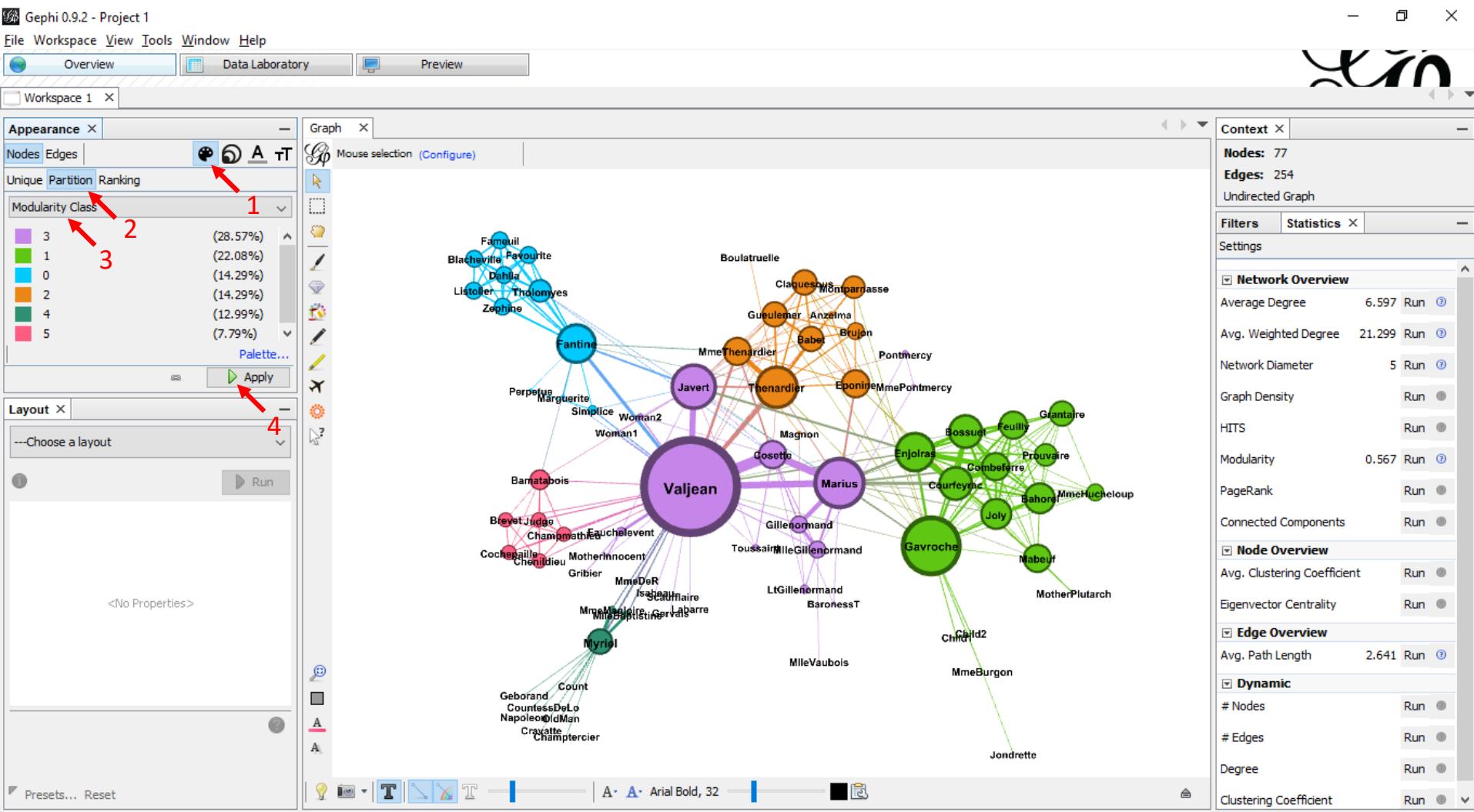
Data Table

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

Id	Label	Interval	Modularity Class	Degree	Weighted Degree	Eccentricity	Closeness Centrality	Harmonic Closeness Centrality	Betweenness Centrality
11	Valjean		3	36	158.0	3.0	0.644068	0.732456	1624.4688
48	Gavroche		1	22	56.0	3.0	0.513514	0.605263	470.570632
55	Marius		3	19	104.0	3.0	0.531469	0.60307	376.292593
27	Javert		3	17	47.0	3.0	0.517007	0.585526	154.844945
25	Thenardier		2	16	61.0	3.0	0.517007	0.58114	213.468481
23	Fantine		0	15	47.0	4.0	0.460606	0.539474	369.486942
58	Enjolras		1	15	91.0	3.0	0.481013	0.552632	121.277067
62	Courfeyrac		1	13	84.0	4.0	0.4	0.483553	15.011035
64	Bossuet		1	13	66.0	3.0	0.475	0.539474	87.647903
63	Bahorel		1	12	39.0	4.0	0.393782	0.472588	6.228642
65	Joly		1	12	43.0	4.0	0.393782	0.472588	6.228642
24	MmeThenardier		2	11	34.0	4.0	0.460606	0.520833	82.656893
26	Cosette		3	11	68.0	4.0	0.477987	0.533991	67.819322
41	Eponine		2	11	19.0	4.0	0.395833	0.470395	32.739519
57	Mabeuf		1	11	16.0	4.0	0.395833	0.470395	78.834524
59	Combeferre		1	11	68.0	4.0	0.391753	0.466009	3.562915
61	Feuilly		1	11	38.0	4.0	0.391753	0.466009	3.562915
0	Myriel		4	10	31.0	4.0	0.429379	0.491228	504.0
66	Grantaire		1	10	16.0	4.0	0.358491	0.432018	0.428571
68	Gueulemer		2	10	25.0	3.0	0.463415	0.517544	14.137094
69	Babet		2	10	27.0	3.0	0.463415	0.517544	14.137094
70	Claquesous		2	10	20.0	4.0	0.452381	0.509868	13.856142
16	Tholomyes		0	9	26.0	4.0	0.391753	0.457237	115.793642
60	Prouvaire		1	9	19.0	4.0	0.356808	0.425439	0.0
71	Montparnasse		2	9	12.0	3.0	0.457831	0.508772	11.040415
29	Bamatabois		5	8	11.0	4.0	0.426966	0.480263	22.916667

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

Color Nodes by Modularity Class



Statics: Connected Components

This example has only one connected component since all nodes are connected.

Try removing some nodes, say “Valjean”, and re-run to see what happens.

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory P HTML Report

Workspace 1 X

Appearance X

Nodes Edges Unique Ranking

Degree

Min size: 1 Max size: 49.5

Spline... |

Graph X

Mouse selection (Config)

Layout X

--Choose a layout

Run

<No Properties>

Presets... Reset

Print Copy Save Close

Connected Components Report

Parameters:
Network Interpretation: undirected

Results:
Number of Weakly Connected Components: 1

Size Distribution

Count

Size (number of nodes)

Context X

Nodes: 77
Edges: 254
Undirected Graph

Filters Statistics

Settings

Network Overview

- Average Degree 6.597 Run
- Avg. Weighted Degree 21.299 Run
- Network Diameter 5 Run
- Graph Density 0.087 Run
- HITS Run
- Modularity 0.566 Run
- PageRank Run
- Connected Components 1 Run

Node Overview

- Avg. Clustering Coefficient Run

Edge Overview

- Avg. Path Length 2.641 Run

Dynamic

- # Nodes Run
- # Edges Run
- Degree Run

HTML Report

Run

Filter Nodes by Degree Range

The figure consists of four panels illustrating the steps to filter nodes by degree range in a network visualization tool.

- Panel 1:** Shows the initial state with the "Topology" node selected in the "Filters" section of the left sidebar. A red arrow labeled "1" points to the "Topology" node.
- Panel 2:** Shows the "Degree Range" query being selected. A red arrow labeled "2" points to the "Degree Range" node under "Queries".
- Panel 3:** Shows the "Degree Range Settings" dialog open. A red arrow labeled "3" points to the "Degree Range" label. The settings show a range from 1 to 36, with the slider set between 10 and 36. A red arrow labeled "4" points to the "Select" button.
- Panel 4:** Shows the final state where the network graph displays only nodes with a degree between 10 and 36. The central node, Valjean, has a degree of 36 and is highlighted in red. Other nodes like Fantine, Javert, and Gavroche also have high degrees and are prominent.

Filter Nodes by Modularity Class

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Appearance X

Nodes Edges Unique Partition Ranking

#c0c0c0

Apply

Layout X

--Choose a layout Run

<No Properties>

Presets... Reset

Graph X

No selection

Context X

Nodes: 77
Edges: 254
Undirected Graph

Filters X Statistics

Reset Add

Library

- Attributes
 - Equal
 - Modularity Class *Intra*(Node)
 - Inter Edges
 - Intra Edges
 - Non-null
 - Partition
 - Partition Count
 - Range
- Dynamic
- Edges
- Operator
- Topology

Queries

Equal (Modularity Class) 1

Equal (Modularity Class) Settings

Value: 8 2

Minimum: 0

Maximum: 8

Select Filter

The screenshot displays the Gephi 0.9.2 software interface for analyzing a network graph. The main workspace shows a complex network structure with nodes colored according to their modularity class. The 'Filters' panel on the right side of the interface is open, showing a hierarchical tree of filter categories. A specific filter, 'Equal (Modularity Class)', has been selected and is highlighted with a red arrow labeled '1'. Below this, the 'Equal (Modularity Class) Settings' dialog is open, showing a 'Value' input field set to '8', with another red arrow labeled '2' pointing to it. The 'Minimum' and 'Maximum' fields are also visible in this dialog. The top menu bar includes options like File, Workspace, View, Tools, Window, and Help. The bottom toolbar contains various tools for node selection and modification.

Layout: Force Atlas (with default parameters)

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1

Appearance X

Nodes Edges Unique Partition Ranking

#c0c0c0

Apply

Layout X

Force Atlas

1

Run

2

Graph X

Dragging (Configure)

Context X

Nodes: 77
Edges: 254
Undirected Graph

Filters Statistics X

Settings

Network Overview

- Average Degree Run
- Avg. Weighted Degree Run
- Network Diameter Run
- Graph Density Run
- HITS Run
- Modularity Run
- PageRank Run
- Connected Components Run

Node Overview

- Avg. Clustering Coefficient Run
- Eigenvector Centrality Run

Edge Overview

- Avg. Path Length Run

Dynamic

- # Nodes Run
- # Edges Run
- Degree Run
- Clustering Coefficient Run

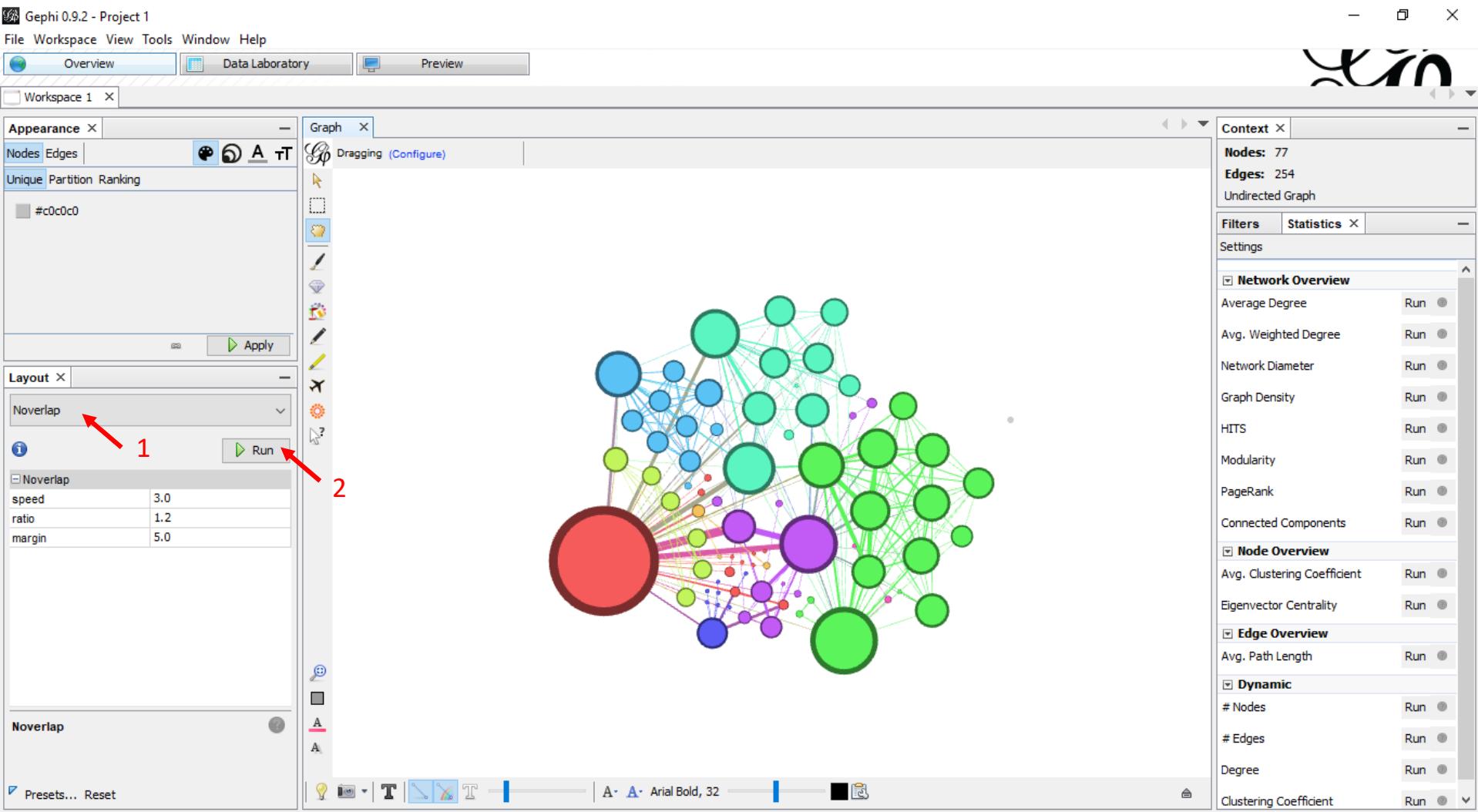
Presets... Reset

Arial Bold, 32

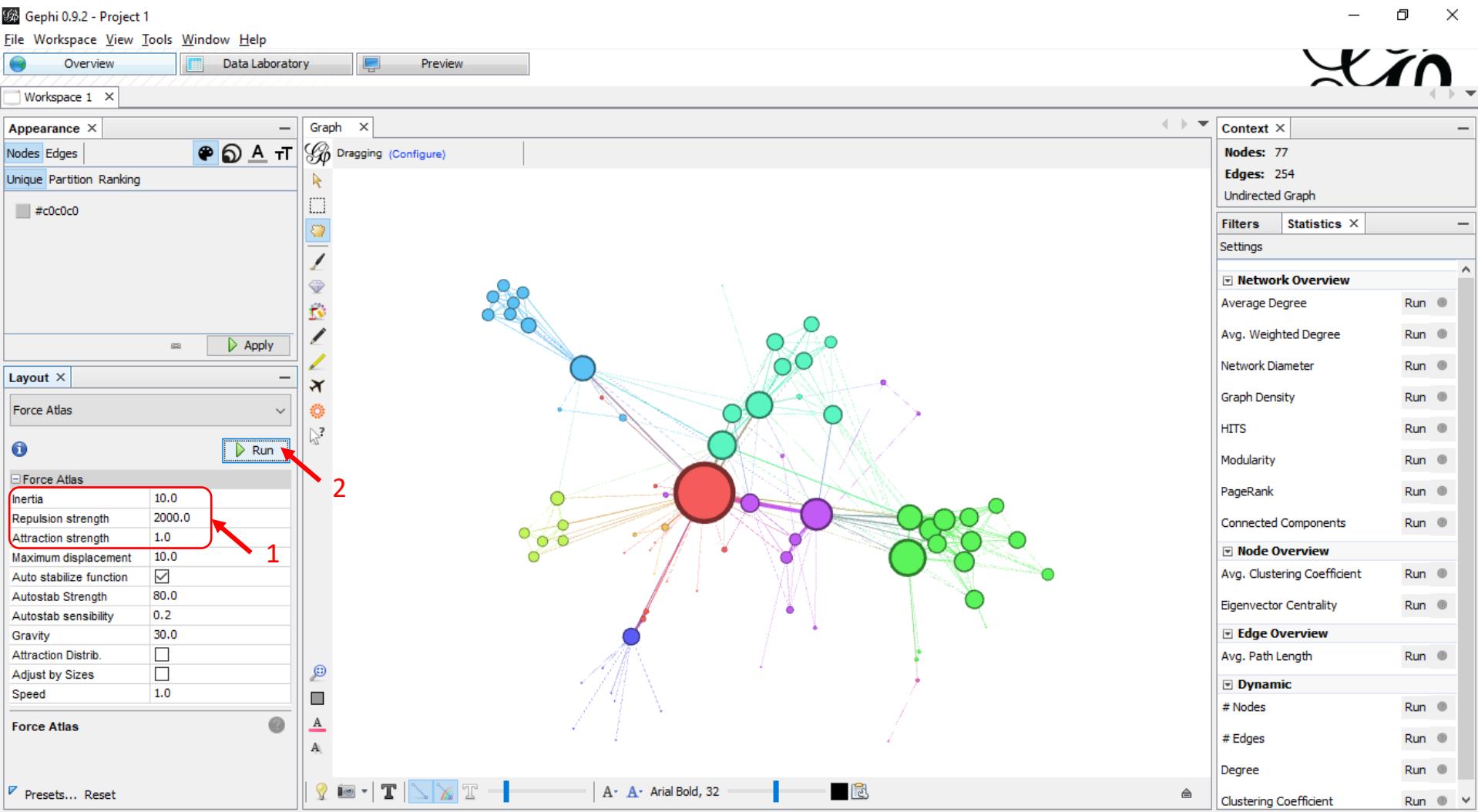
Color palette

Tool icons: Selection, Zoom, Edge, Node, Text, Shape, Filter, Layout, Statistics, Network, Node, Edge, Dynamic, Presets.

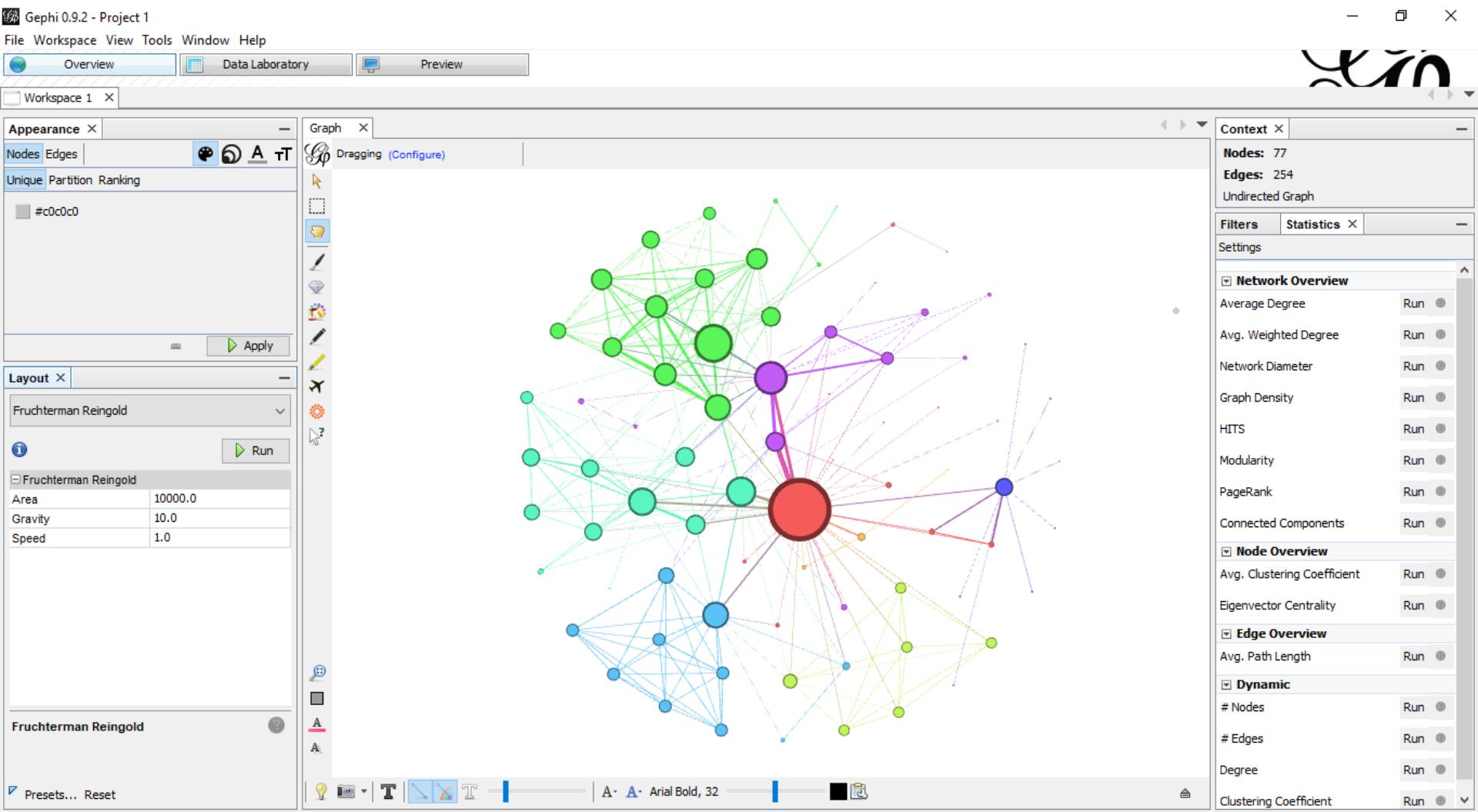
Layout: Force Atlas (with default parameters) + Noverlap



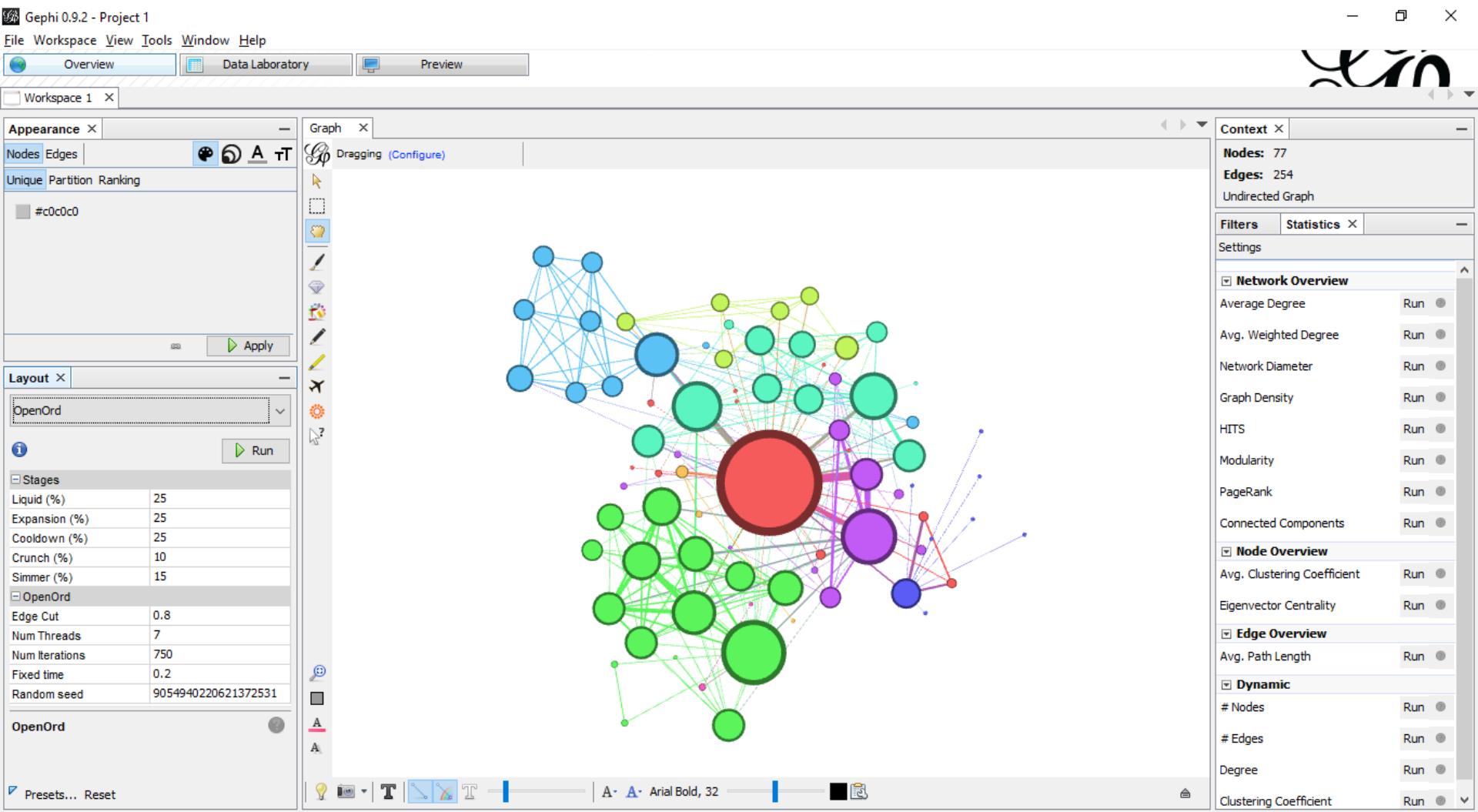
Layout: Force Atlas (with adjusted parameters)



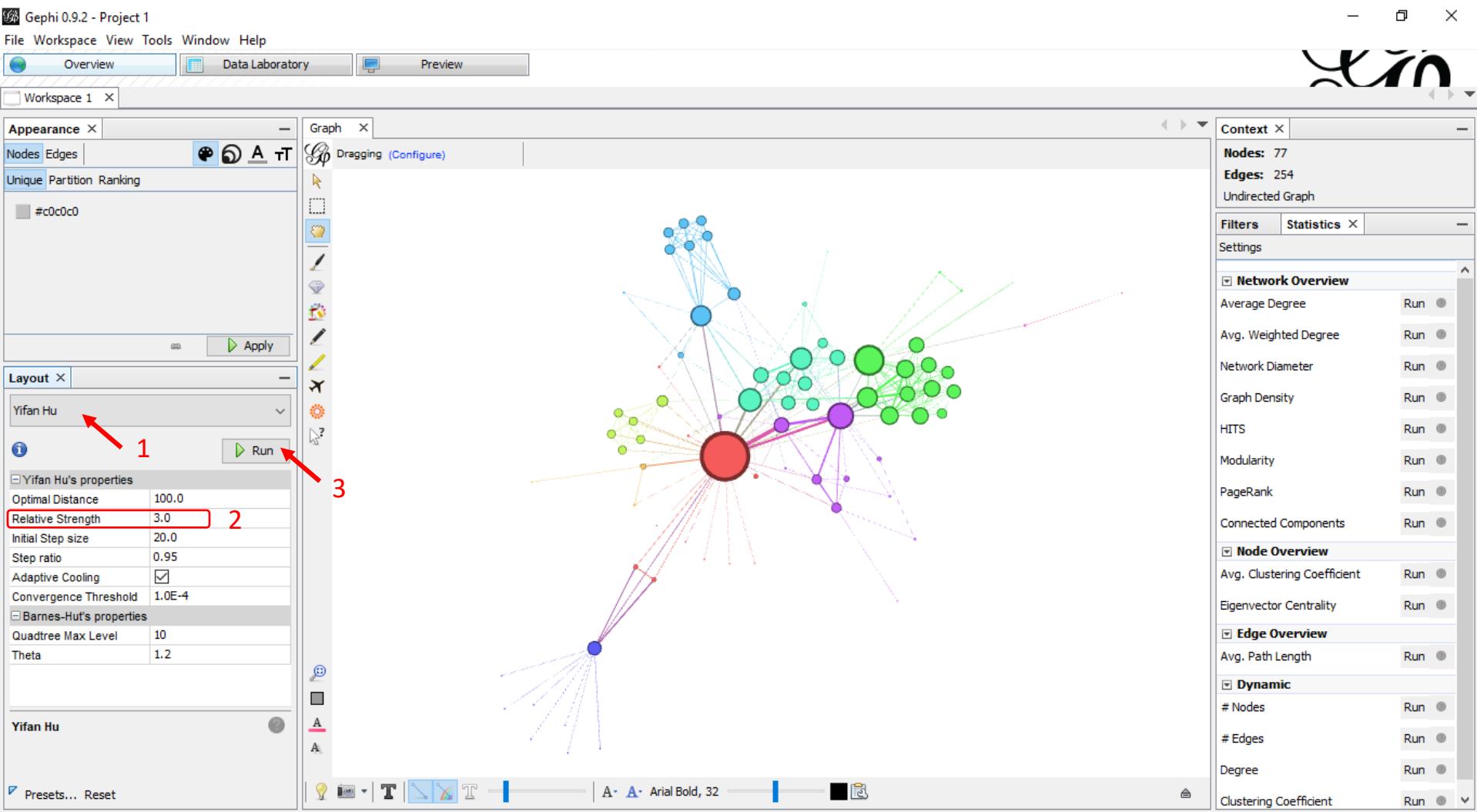
Layout: Fruchterman Reingold



Layout: OpenOrd + Noverlap



Layout: Yifan Hu



Data Preparation for Gephi

- <http://humnviz.blogs.bucknell.edu/files/2015/11/Data-Preparation-for-Gephi.pdf>
- <https://github.com/seinecle/gephi-tutorials/blob/master/src/main/asciidoc/en/importing-csv-data-in-gephi-en.adoc>

Gephi File Types

- Gephi project file format
 - gephi
- Some supported network file formats
 - gexf csv graphml gdf gml

Sample Network Datasets

- Gephi sample dataset
<https://github.com/gephi/gephi/wiki/Datasets>
- Stanford Large Network Dataset Collection (really large)
<http://snap.stanford.edu/data/>

Importing Network Data

Method 1: only node and edge

- A single csv file with two columns for source and target

Source,Target

Bangkok,Chonburi

Phuket, Chiangmai

Method 2: node, edge, and attributes

- Two files for nodes and edges
- Node file

Id,Label,Date of Birth,Place of Birth

1,"Jeremy,Jennifer",25/03/1978,"Tampa"

45,"Rodriguez, Valerian",30/04/1985,"Berlin"

- Edge file

Source,Target,Weight,Type,Where first met

1,45,3,"undirected",London

Import a CSV File

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

New Project Ctrl+Shift+N
Open... Ctrl+O
Open Recent...
Close Project
Properties...

Import spreadsheet...

Import Database
Import...
Generate

Save Ctrl+S
Save As...

Export

Exit

Layout X
--Choose a layout
<No Properties>
Run
Presets... Reset

Preview
Workspace 3 × Workspace 4 × Workspace 5 ×

Mouse selection (Configure)

Context X
Nodes: 0
Edges: 0
Directed Graph

Filters Statistics X
Settings

Network Overview
Average Degree Run
Avg. Weighted Degree Run
Network Diameter Run
Graph Density Run
HITS Run
Modularity Run
PageRank Run
Connected Components Run

Node Overview
Avg. Clustering Coefficient Run
Eigenvector Centrality Run

Edge Overview
Avg. Path Length Run

Dynamic
Nodes Run

Timeline disabled. The graph is not dynamic.

The screenshot shows the Gephi 0.9.2 application window. The title bar reads "Gephi 0.9.2 - Project 1". The "File" menu is open, displaying various options like "New Project", "Open...", and "Import spreadsheet...". The "Import spreadsheet..." option is highlighted with a blue selection bar. The main workspace is currently empty. On the right side, there are several panels: "Context" (Nodes: 0, Edges: 0, Directed Graph), "Network Overview" (with buttons for Average Degree, Avg. Weighted Degree, Network Diameter, Graph Density, HITS, Modularity, PageRank, and Connected Components), "Node Overview" (Avg. Clustering Coefficient, Eigenvector Centrality), "Edge Overview" (Avg. Path Length), and "Dynamic" (# Nodes). At the bottom, a status bar says "Timeline disabled. The graph is not dynamic.".

retweet.csv

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1 × Workspace 2 × Workspace 3 × Workspace 4 × Workspace 5 ×

Appearance X

Nodes Edges

Unique Partition Ranking

#c0c0c0

Graph X

Mouse selection (Continuous)

Spreadsheet (CSV)...

Steps

1. General CSV options
2. Import settings

General CSV options (1 of 2)

CSV file to import:
C:\Users\Pom\Documents\1 Courses\Data Sci 446\Data\retweet.csv

Separator: Import as: Charset:

Comma Adjacency list UTF-8

Preview:

timesofindia	shramm
timesofindia	PoojaParvati
timesofindia	rohitagarwal86
timesofindia	AnandMadh...
timesofindia	prateektan...
timesofindia	ricketyship
timesofindia	surinder0
timesofindia	nehasanghvi
timesofindia	CarieIsSoVery

< Back Next > Finish Cancel Help

Context X

Nodes: 0
Edges: 0
Directed Graph

Filters Statistics X

Settings

Network Overview

- Average Degree Run
- Avg. Weighted Degree Run
- Network Diameter Run
- Graph Density Run
- HITS Run
- Modularity Run
- PageRank Run
- Connected Components Run

Node Overview

- Avg. Clustering Coefficient Run
- Eigenvector Centrality Run

Edge Overview

- Avg. Path Length Run

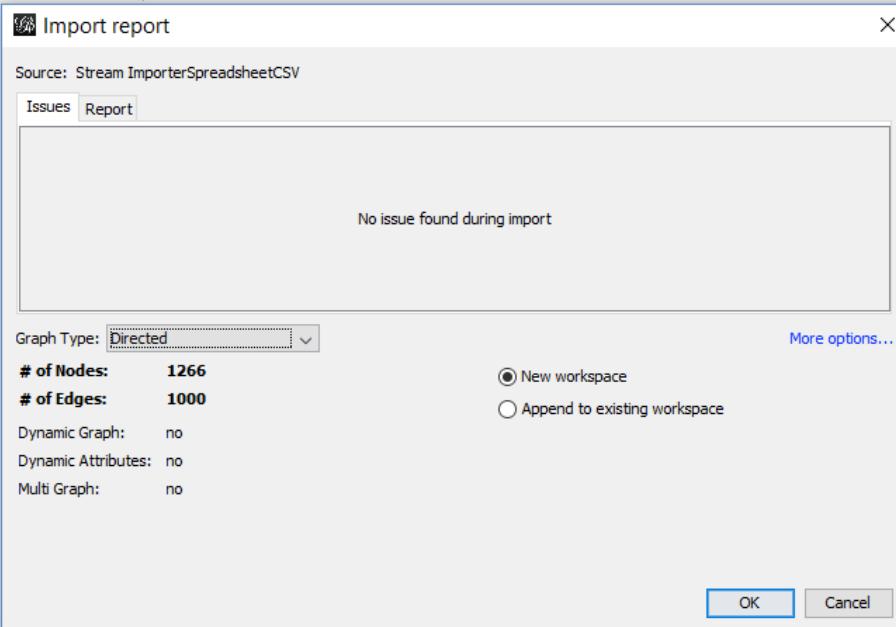
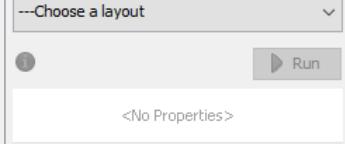
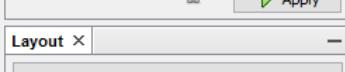
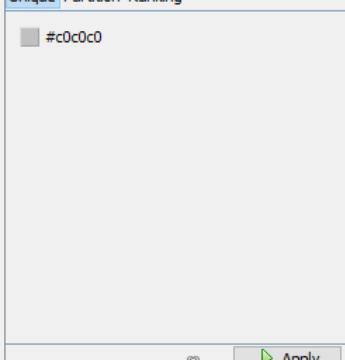
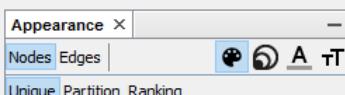
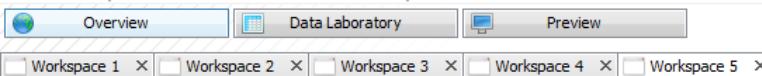
Dynamic

- # Nodes Run

Timeline disabled. The graph is not dynamic.

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help



Context X

Nodes: 0

Edges: 0

Directed Graph

Filters Statistics X

Settings

Network Overview

Average Degree Run

Avg. Weighted Degree Run

Network Diameter Run

Graph Density Run

HITS Run

Modularity Run

PageRank Run

Connected Components Run

Node Overview

Avg. Clustering Coefficient Run

Eigenvector Centrality Run

Edge Overview

Avg. Path Length Run

Dynamic

Nodes Run

Timeline disabled. The graph is not dynamic.

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1 X Workspace 2 X Workspace 3 X Workspace 4 X Workspace 5 X Workspace 7 X



Appearance X

Graph X

Mouse selection (Configure)

Nodes Edges



Unique Partition Ranking



Apply

Layout X

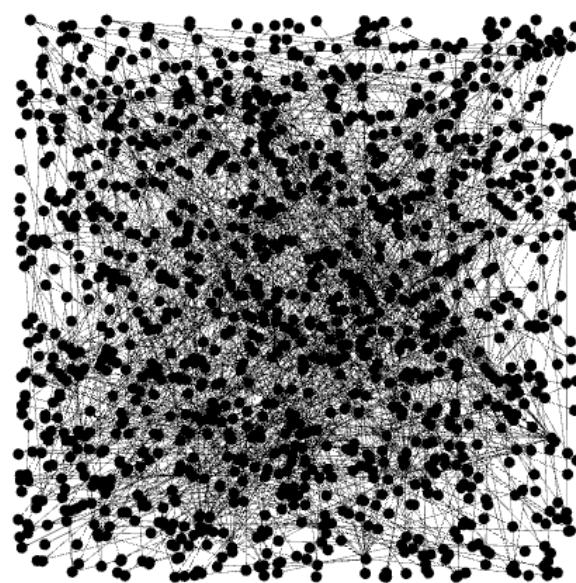
--Choose a layout

Run

<No Properties>



Presets... Reset



Context X

Nodes: 1266

Edges: 1000

Directed Graph

Filters Statistics

Settings

Network Overview

Average Degree

Run

Avg. Weighted Degree

Run

Network Diameter

Run

Graph Density

Run

HITS

Run

Modularity

Run

PageRank

Run

Connected Components

Run

Node Overview

Avg. Clustering Coefficient

Run

Eigenvector Centrality

Run

Edge Overview

Avg. Path Length

Run

Dynamic

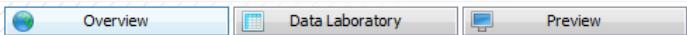
Nodes

Run

Timeline disabled. The graph is not dynamic.

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help



Workspace 1 X Workspace 2 X Workspace 3 X Workspace 4 X Workspace 5 X Workspace 7 X



Appearance X

Nodes Edges

Unique Ranking

Size: 3

Apply

Layout X

ForceAtlas 2

Run

Performance

Tolerance (speed) 10.0

Approximate Repulsion

Approximation 1.2

Tuning

Scaling 2.0

Stronger Gravity

Gravity 15.0

Behavior Alternatives

Dissuade Hubs

LinLog mode

Prevent Overlap

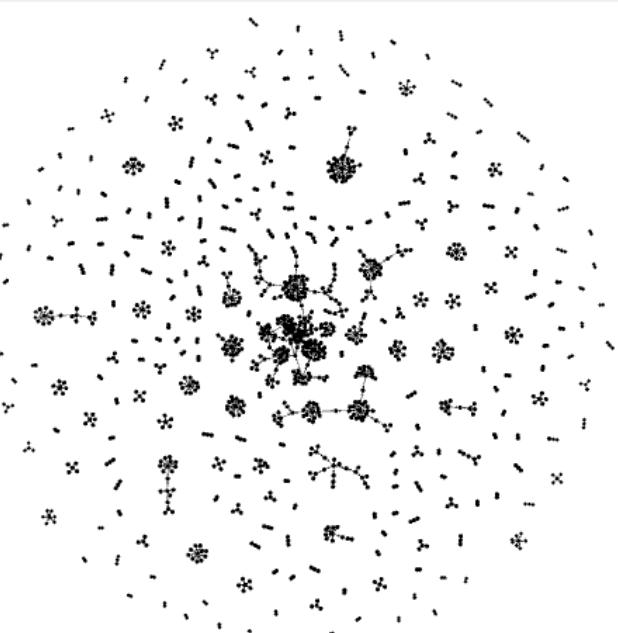
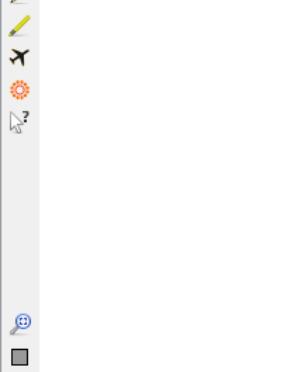
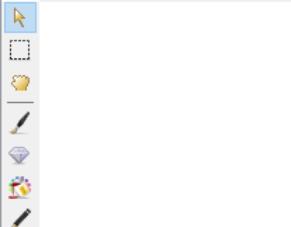
Edge Weight Influence 1.0

ForceAtlas 2

Presets... Reset

Graph X

Mouse selection (Configure)



Context X

Nodes: 1266

Edges: 1000

Directed Graph

Filters Statistics X

Settings

Network Overview

Average Degree Run

Avg. Weighted Degree Run

Network Diameter 2 Run

Graph Density Run

HITS Run

Modularity Run

PageRank Run

Connected Components Run

Node Overview

Avg. Clustering Coefficient Run

Eigenvector Centrality Run

Edge Overview

Avg. Path Length 1.045 Run

Dynamic

Nodes Run

Timeline disabled. The graph is not dynamic.

Filter by Giant Component

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1 × Workspace 2 × Workspace 3 × Workspace 4 × Workspace 5 × Workspace 7 ×

Appearance X

Nodes Edges Unique Ranking

Size: 3 ↕ Apply

Layout X

ForceAtlas 2

Tolerance (speed) 10.0
Approximate Repulsion
Approximation 1.2
Tuning
Scaling 2.0
Stronger Gravity
Gravity 15.0
Behavior Alternatives
Dissuade Hubs
LinLog mode
Prevent Overlap
Edge Weight Influence 1.0

ForceAtlas 2

Presets... Reset

Graph X

No selection

Context X

Nodes: 1266
Edges: 1000
Directed Graph

Filters X Statistics

Library

- Attributes
- Dynamic
- Edges
- Operator
- Topology
 - Degree Range
 - Ego Network
 - Giant Component
 - Has Self-loop
 - In Degree Range
 - K-core
 - Mutual Degree Range
 - Neighbors Network
 - Out Degree Range

Queries

- + Giant Component

Select Filter

Timeline disabled. The graph is not dynamic.

The screenshot illustrates the process of identifying the largest connected component in a network using Gephi's filtering capabilities. The 'Topology' section of the 'Filters' panel is expanded, and the 'Giant Component' option is selected, as indicated by the red circles. This filter is then applied to the graph, which is visualized using the ForceAtlas 2 layout algorithm.

Gephi 0.9.2 - Project 1

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Workspace 1 X Workspace 2 X Workspace 3 X Workspace 4 X Workspace 5 X Workspace 7 X



Appearance X
Nodes Edges
Unique Ranking



Size: 3

Apply

Layout X

ForceAtlas 2

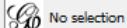
Run

Performance
Tolerance (speed) 10.0
Approximate Repulsion
Approximation 1.2
Tuning
Scaling 2.0
Stronger Gravity
Gravity 15.0
Behavior Alternatives
Dissuade Hubs
LinLog mode
Prevent Overlap
Edge Weight Influence 1.0

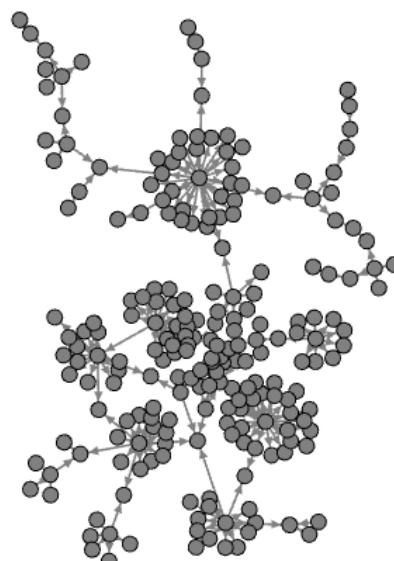
ForceAtlas 2

Presets... Reset

Graph X



No selection



Context X

Nodes: 220 (17.38% visible)

Edges: 220 (22% visible)

Directed Graph

Filters X Statistics

Reset

Library

- Attributes
- Dynamic
- Edges
- Operator
- Topology
 - Degree Range
 - Ego Network
 - Giant Component
 - Has Self-loop
 - In Degree Range
 - K-core
 - Mutual Degree Range
 - Neighbors Network
 - Out Degree Range

Queries

- Giant Component

Select

Stop

Timeline disabled. The graph is not dynamic.

Color by Modularity Class

