

Networks in urbanism and architecture: lecture 2

L.Bauer lecture
LPI, Paris, France
SEA Mumbai, India

Course overview

1. Introduction to networks, urbanism and architecture
Quantitative and qualitative measures for networks
2. Analysis of city systems
3. Projects discussions

Structure of each lecture

theory

practice

Course topics

1. Networks theory and data science
How to quantitatively characterise networks?
2. Examples: smart and non smart cities data
3. City indices: how to quantify cities?

theory

4. Discussions about projects
Colab for loading street networks of Mumbai

practice

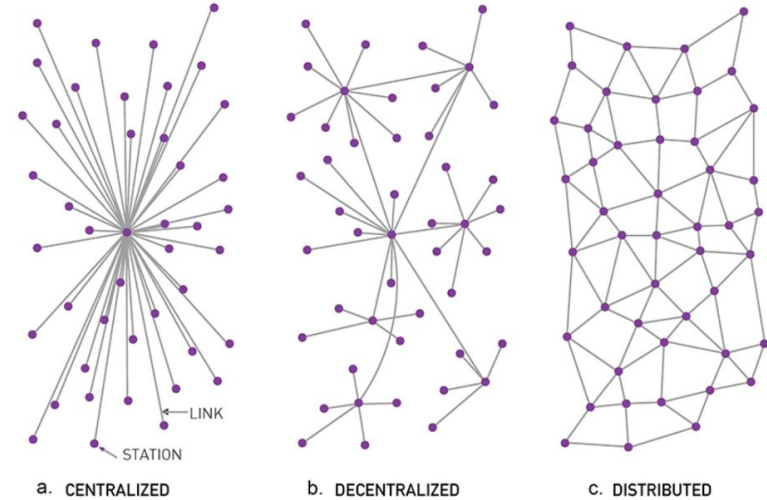
Part 1

Networks and data science

What we will look at in network science?

1. **Network definition and measures**
2. Networks in time and space
3. Networks from data

Fig. credits P. Barran.



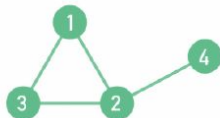
1. Network definitions

How to define a network? From real system to data

a. Adjacency matrix

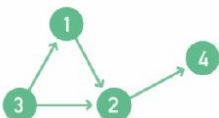
$$A_{ij} = \begin{matrix} & A_{11} & A_{12} & A_{13} & A_{14} \\ & A_{21} & A_{22} & A_{23} & A_{24} \\ & A_{31} & A_{32} & A_{33} & A_{34} \\ & A_{41} & A_{42} & A_{43} & A_{44} \end{matrix}$$

b. Undirected network



$$A_{ij} = \begin{matrix} & 0 & 1 & 1 & 0 \\ & 1 & 0 & 1 & 1 \\ & 1 & 1 & 0 & 0 \\ & 0 & 1 & 0 & 0 \end{matrix}$$

c. Directed network



$$A_{ij} = \begin{matrix} & 0 & 0 & 1 & 0 \\ & 1 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 0 \\ & 0 & 1 & 0 & 0 \end{matrix}$$



1. Network measures and definitions: density



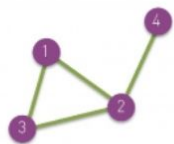
low



high

1. Network types

a. Undirected

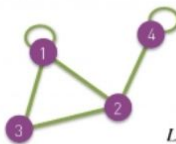


$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$A_{ii} = 0 \quad A_{ij} = A_{ji}$$

$$L = \frac{1}{2} \sum_{i,j=1}^N A_{ij} \quad \langle k \rangle = \frac{2L}{N}$$

b. Self-loops

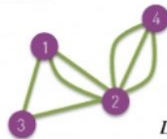


$$A_{ij} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

$$\exists i, A_{ii} \neq 0 \quad A_{ij} = A_{ji}$$

$$L = \frac{1}{2} \sum_{i,j=1}^N A_{ij} + \sum_{i=1}^N A_{ii} \quad ?$$

c. Multigraph
(undirected)

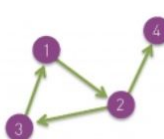


$$A_{ij} = \begin{pmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$A_{ii} = 0 \quad A_{ij} = A_{ji}$$

$$L = \frac{1}{2} \sum_{i,j=1}^N A_{ij} \quad \langle k \rangle = \frac{2L}{N}$$

d. Directed

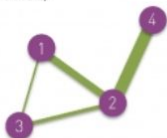


$$A_{ij} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A_{ij} \neq A_{ji}$$

$$L = \sum_{i,j=1}^N A_{ij} \quad \langle k \rangle = \frac{L}{N}$$

e. Weighted
(undirected)

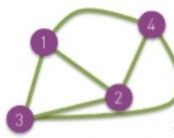


$$A_{ij} = \begin{pmatrix} 0 & 2 & 0.5 & 0 \\ 2 & 0 & 1 & 4 \\ 0.5 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \end{pmatrix}$$

$$A_{ii} = 0 \quad A_{ij} = A_{ji}$$

$$\langle k \rangle = \frac{2L}{N}$$

f. Complete Graph
(undirected)

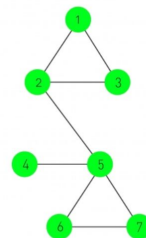


$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

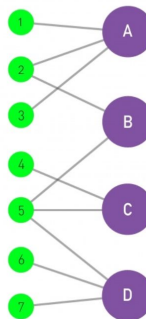
$$A_{ii} = 0 \quad A_{ij} = 1$$

$$L = L_{\max} = \frac{N(N-1)}{2} \quad \langle k \rangle = N-1$$

PROJECTION U U

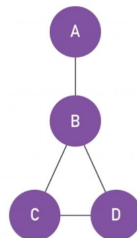


U

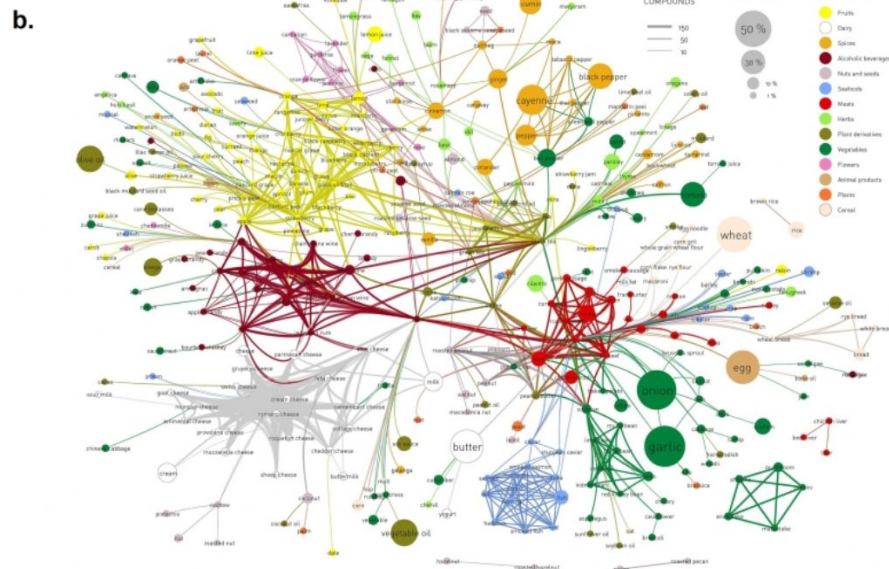
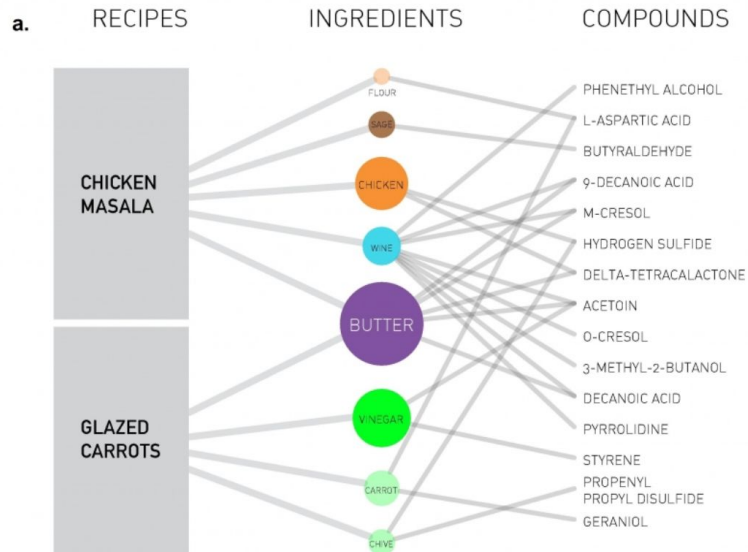


V

PROJECTION V V



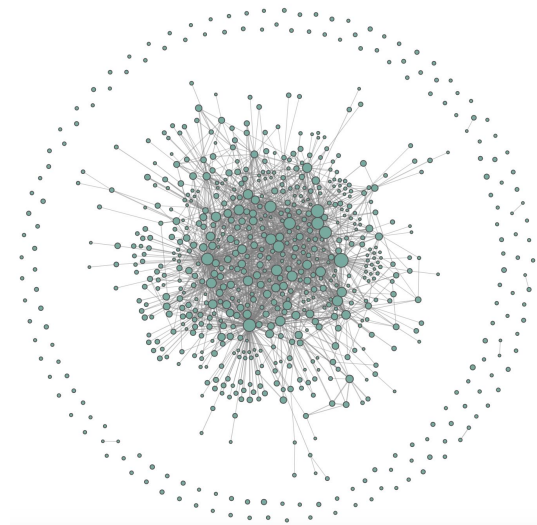
Network types: bipartite networks



1. Network measures:

Local measures for each node

Global measures for the whole network



Network measures: Check calculation on networkx python library

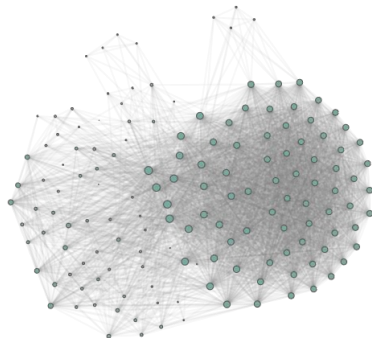
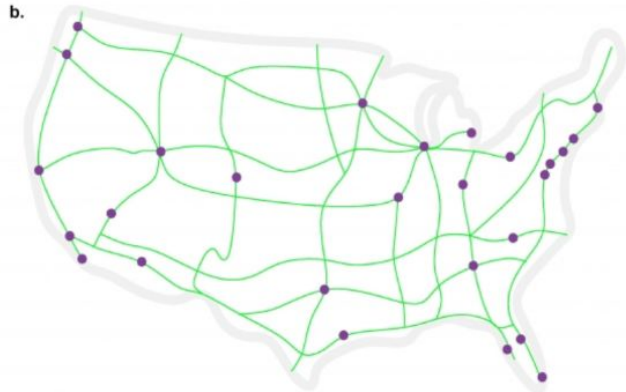


TABLE 2: Definitions of network science terms and variables.

Term/variable	Definition
N	number of nodes, N , in graph
E	number of edges, E , in graph
network density	ratio of the number of edges to the maximum number of possible edges $\frac{2E}{N(N-1)}$
distance, $d(n_i, n_j)$	shortest path between node i and node j $d(n_i, n_j)$ where $n_i, n_j \in N$
average shortest path length, L	average length of shortest path between pairs of nodes $L = \frac{1}{N(N-1)} \cdot \sum_{i \neq j} d(n_i, n_j)$
diameter, D	largest shortest path between nodes $D = \max_{n_i \in N, n_j \in N} d(n_i, n_j)$
closeness centrality	inverse of the sum of the length of the shortest paths between node i and all other nodes in the graph $C_i = \frac{1}{\sum_j d(n_i, n_j)}$
degree, k_i	number of edges attached to node i
average degree, $\langle k \rangle$	average number of edges per node in network $\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i$
local clustering coefficient, g_i	number of edges between the neighbors of node i divided by the maximum number of edges between those neighbors $g_i = \frac{2 e_{jk} }{k_i(k_i-1)} \text{ where } n_j, n_k \in N_i, e_{jk} \in E$
average clustering coefficient, $\langle C \rangle$	average clustering coefficient of nodes in the network $\langle C \rangle = \frac{1}{N} \sum_{i=1}^N g_i$
modularity, Q	proportion of edges that fall within subgroups of nodes minus the expected proportion if edges were randomly distributed, range $[-1, 1]$
average efficiency, E_G	measure of how efficiently information is exchanged in the network $E_G = \frac{1}{n(n-1)} \sum_{i \neq j \in N} \frac{1}{d(n_i, n_j)}$
largest connected component	largest group of nodes in the network that are connected to each other in a single component
degree distribution, $P(k)$	probability distribution of node degrees in the network
γ	power-law exponent for the degree distribution
Small world structure	network with short average path lengths and relatively high clustering coefficient (relative to a random graph with similar density)
scale-free network	network with a degree distribution that is power-law distributed

1. Network measures:

Degree measure look more into <http://networksciencebook.com/chapter/2>

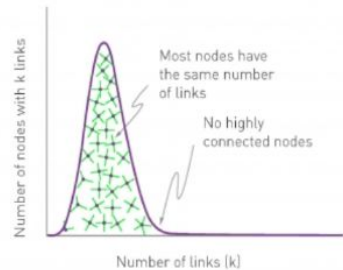


1. Network measures:

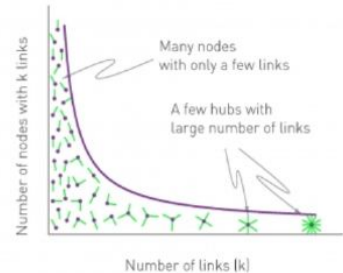
Degree or why do we care? How to look into degree distributions?



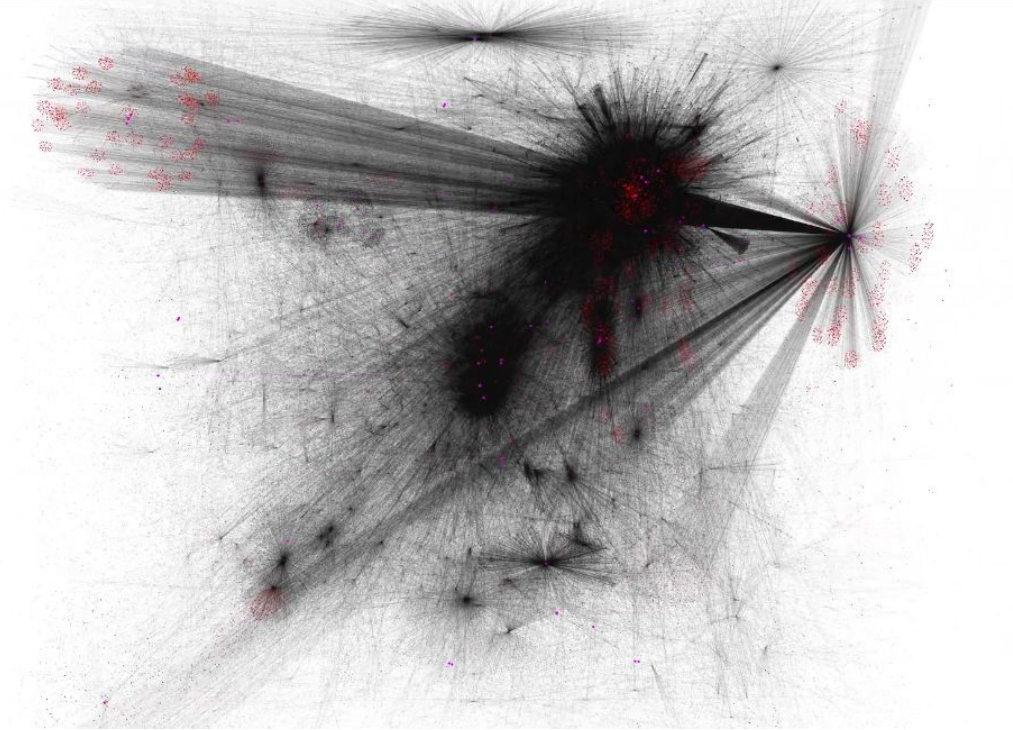
a. POISSON



c. POWER LAW



Visualisation resource



<https://gephi.org/users/download/>

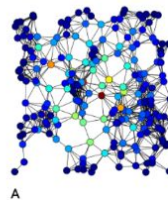
Network measures:

Check calculation on networkx python library

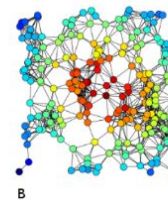
TABLE 2: Definitions of network science terms and variables.

Term/variable	Definition
N	number of nodes, N , in graph
E	number of edges, E , in graph
network density	ratio of the number of edges to the maximum number of possible edges $\frac{2E}{N(N-1)}$
distance, $d(n_i, n_j)$	shortest path between node i and node j $d(n_i, n_j)$ where $n_i, n_j \in N$
average shortest path length, L	average length of shortest path between pairs of nodes $L = \frac{1}{N(N-1)} \sum_{i \neq j} d(n_i, n_j)$
diameter, D	largest shortest path between nodes $D = \max_{n_i, n_j \in N} d(n_i, n_j)$
closeness centrality	inverse of the sum of the length of the shortest paths between node i and all other nodes in the graph $C_i = \frac{1}{\sum_j d(n_i, n_j)}$
degree, k_i	number of edges attached to node i
average degree, $\langle k \rangle$	average number of edges per node in network $\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i$
local clustering coefficient, c_i	number of edges between the neighbors of node i divided by the maximum number of edges between those neighbors $c_i = \frac{2 e_{\mu\lambda} }{k_i(k_i-1)}$ where $n_\mu, n_\lambda \in N$, $e_{\mu\lambda} \in E$
average clustering coefficient, $\langle C \rangle$	average clustering coefficient of nodes in the network $\langle C \rangle = \frac{1}{N} \sum_{i=1}^N c_i$
modularity, Q	proportion of edges that fall within subgroups of nodes minus the expected proportion if edges were randomly distributed, range $[-1, 1]$
average efficiency, E_G	measure of how efficiently information is exchanged in the network $E_G = \frac{1}{n(n-1)} \sum_{i \neq j \in N} \frac{1}{d(n_i, n_j)}$
largest connected component	largest group of nodes in the network that are connected to each other in a single component
degree distribution, $P(k)$	probability distribution of node degrees in the network
γ	power-law exponent for the degree distribution
Small world structure	network with short average path lengths and relatively high clustering coefficient (relative to a random graph with similar density)
scale-free network	network with a degree distribution that is power-law distributed

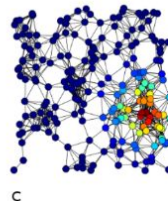
Betweenness centrality



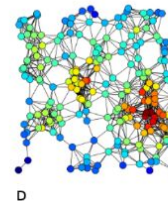
Closeness centrality



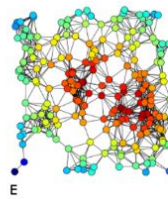
Eigenvector centrality



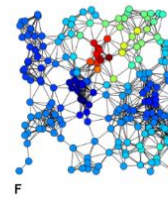
Degree centrality



Harmonic centrality



Katz centrality

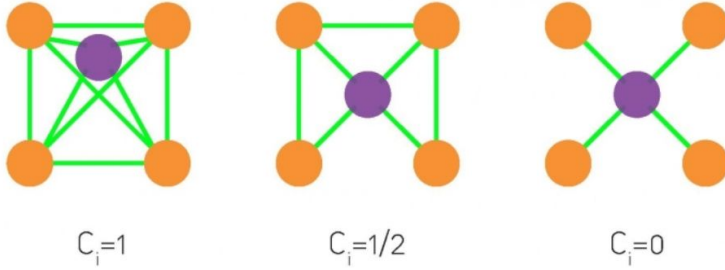


1. Network measures

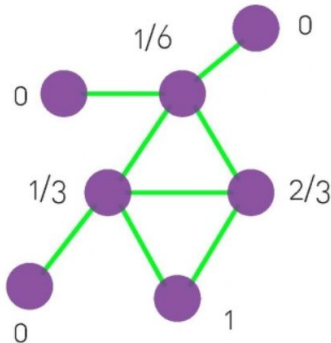
Example of clustering

Notebooks at https://github.com/Big-data-course-CRI/materials_big_data_cri_2019

a.



b.



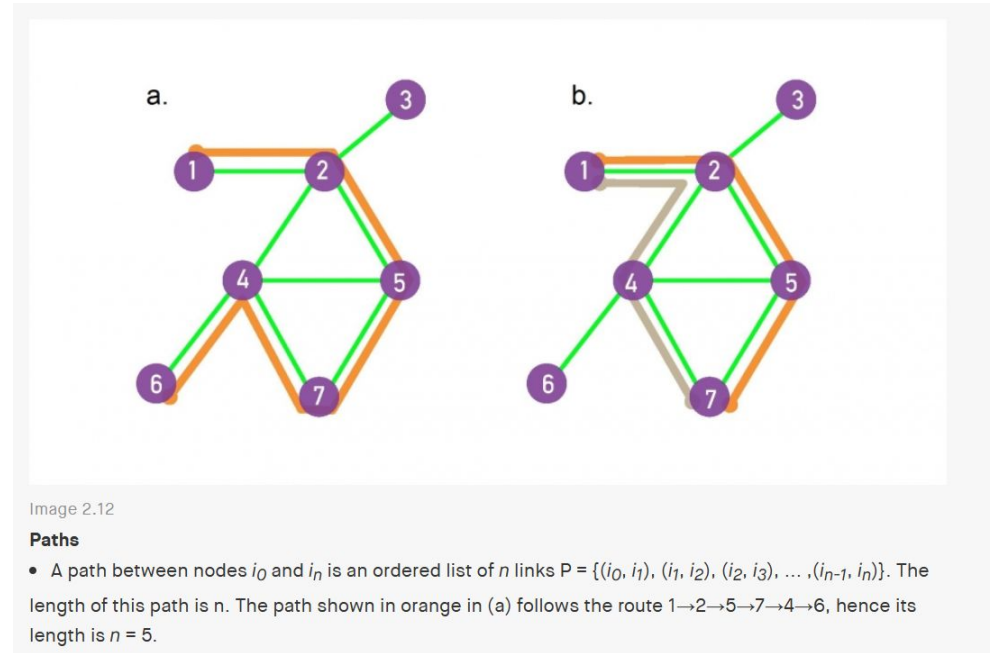
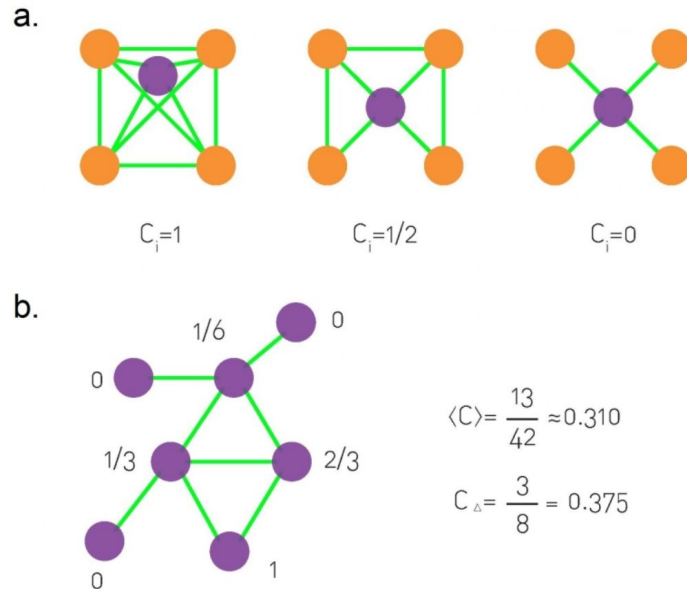
$$\langle C \rangle = \frac{13}{42} \approx 0.310$$

$$C_{\Delta} = \frac{3}{8} = 0.375$$

1. Network measures

Example of clustering, paths, betweenness

Notebooks at https://github.com/Big-data-course-CRI/materials_big_data_cri_2019



What we will look at in network science?

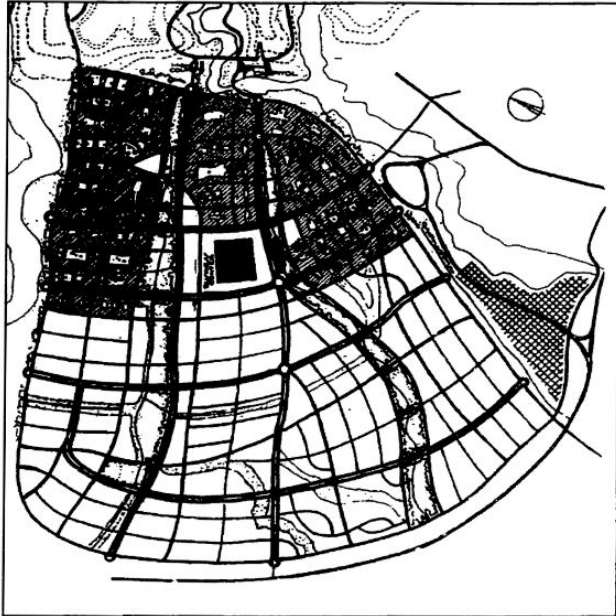
1. Network measures
2. **Networks in time and space**
3. Networks from data

Figure 7.11

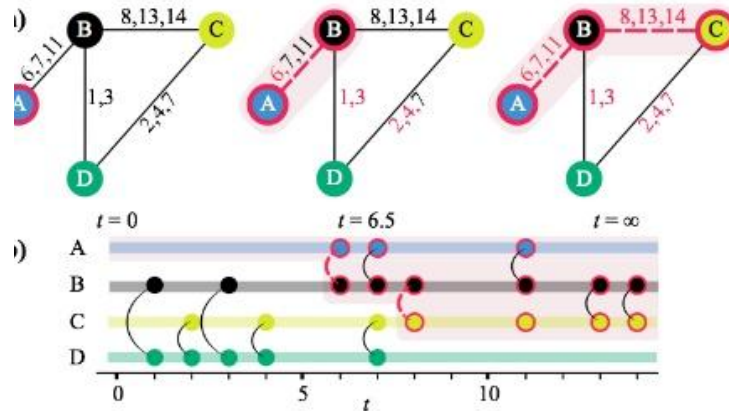
Aaron Koblin's *Flight Patterns* (2005): visualization of the flight paths of aircraft crossing North America

2. Networks in time and space

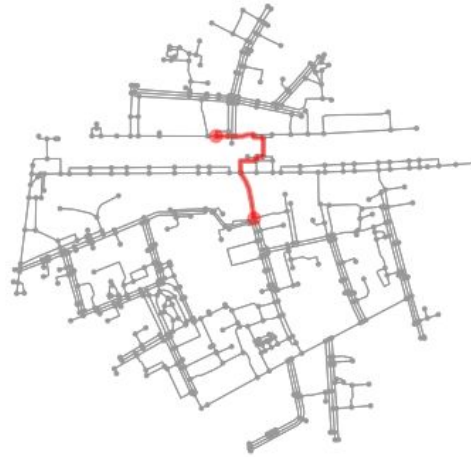
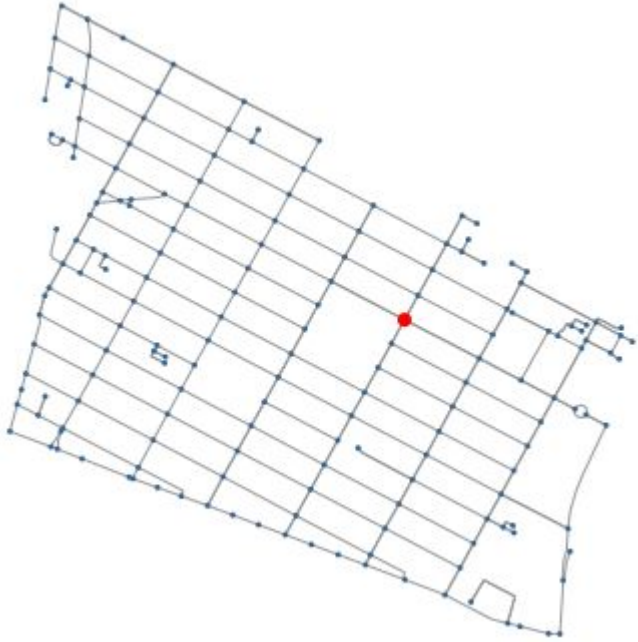
Master Plan for Chandigarh by Albert Mayer RAIC Journal, 1955 (Evenson Norma, Chandigarh, 1966)



Temporality matters:
reachability issue



2. Networks in time and space

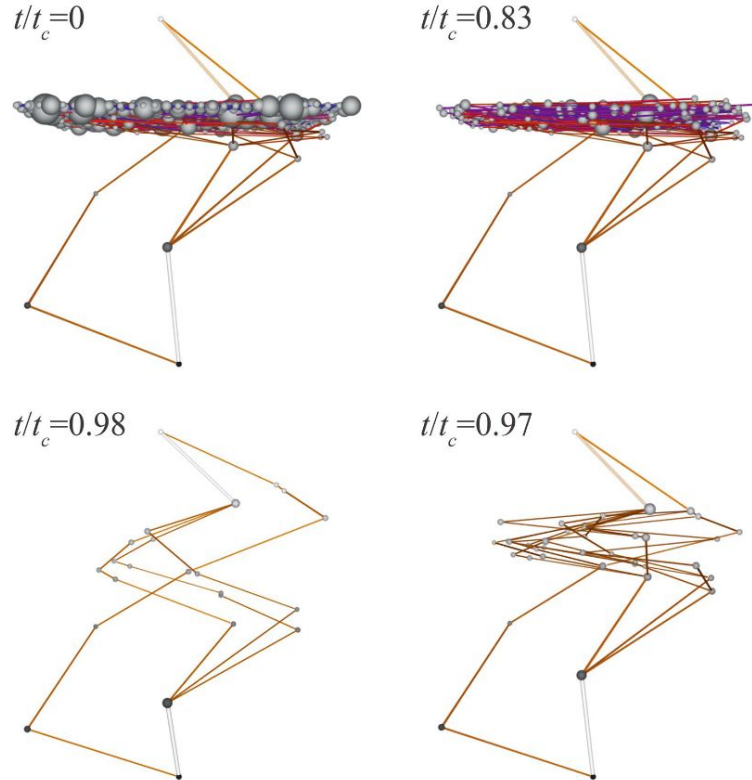


Osmnx for spatial networks
analysis

<https://arxiv.org/abs/1010.0302>

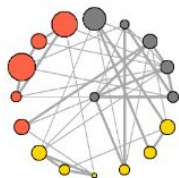
Networks in time and space

<https://www.nature.com/articles/s41598-019-44701-6>

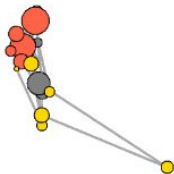


Networks layout (also for Day 3)

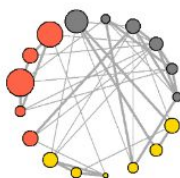
layout_as_star



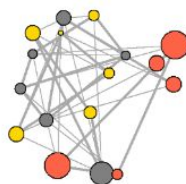
layout_components



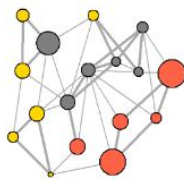
layout_in_circle



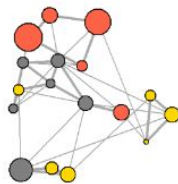
layout_randomly



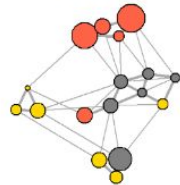
layout_with_dh



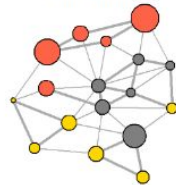
layout_with_drl



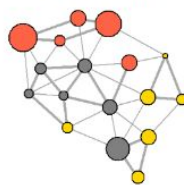
layout_with_fr



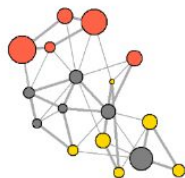
layout_with_gem



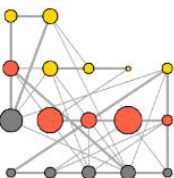
layout_with_graphopt



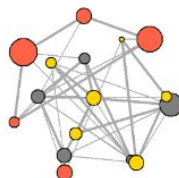
layout_nicely



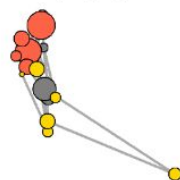
layout_on_grid



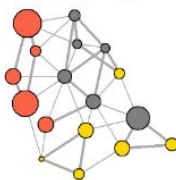
layout_on_sphere



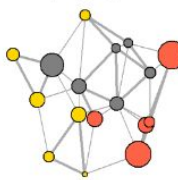
layout_with_kk



layout_with_lgl



layout_with_mds



Where can I get network data?

Example:

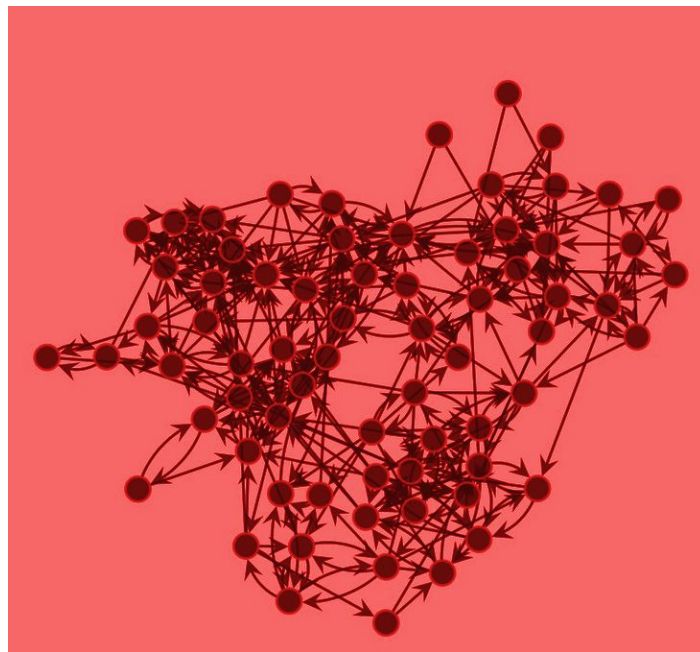
Highschool: Illinois high school students (1958). A network of friendships among male students in a small high school in Illinois from 1958. 70 nodes, 366 edges.

<https://networks.skewed.de/net/highschool>

Example:

Facebook or wikipedia data

<https://snap.stanford.edu/data/wiki-meta.html>



What we will look at in network science?

1. Network measures
2. Networks in time and space
3. **Networks from data**

Figure 7.11

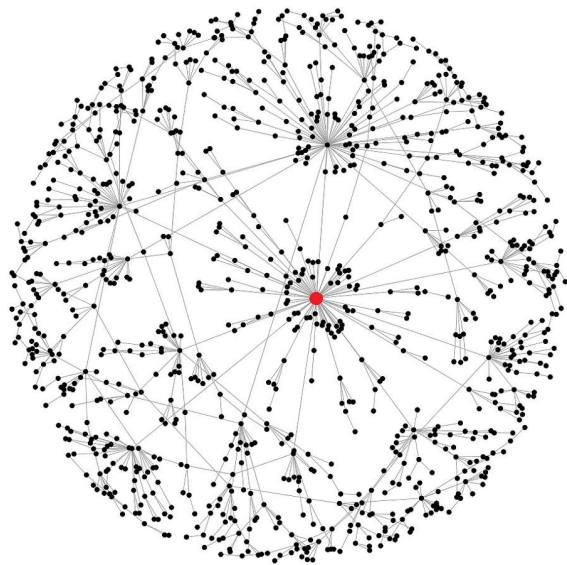
Aaron Koblin's *Flight Patterns* (2005): visualization of the flight paths of aircraft crossing North America

How to construct networks from data?

1. Directly build correspondence between links and edges with data (social networks, flights data)
2. Preprocess data (first build correlation from data)

How to construct networks from data?

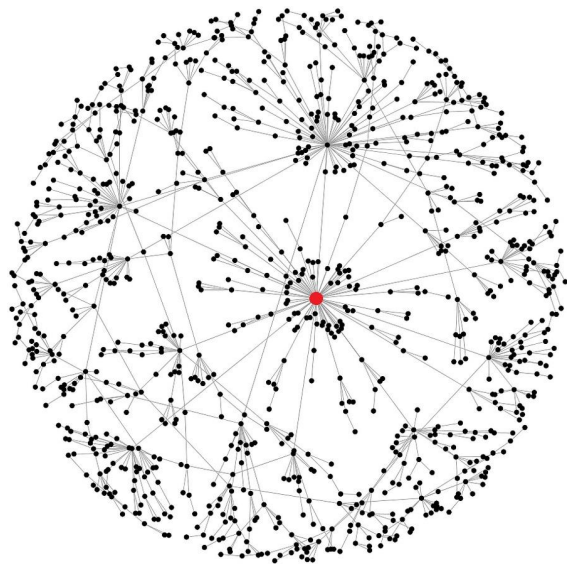
1. Directly build correspondence between links and edges with data (social networks, flights data)
2. Preprocess data (first build correlation from data



How to construct networks from data?

1. Directly build correspondence between links and edges with data (social networks, flights data)
2. Preprocess data (first build correlation from data

Remember that networks do not give one-to-one
Correspondence of your data.
Hence do not generalize



Social networks analysis

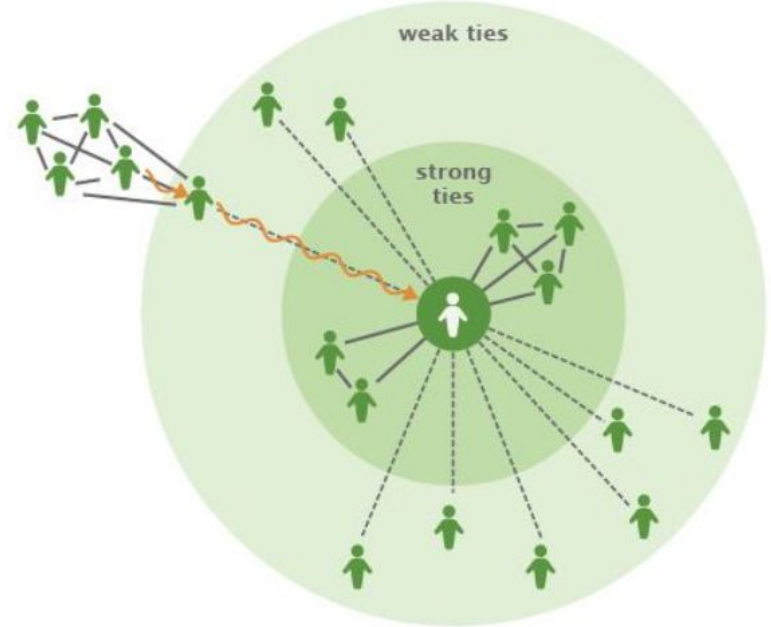
The Strength of Weak Ties¹

Mark S. Granovetter
Johns Hopkins University

Analysis of social networks is suggested as a tool for linking micro and macro levels of sociological theory. The procedure is illustrated by elaboration of the macro implications of one aspect of small-scale interaction: the strength of dyadic ties. It is argued that the degree of overlap of two individuals' friendship networks varies directly with the strength of their tie to one another. The impact of this principle on diffusion of influence and information, mobility opportunity, and community organization is explored. Stress is laid on the cohesive power of weak ties. Most network models deal, implicitly, with strong ties, thus confining their applicability to small, well-defined groups. Emphasis on weak ties lends itself to discussion of relations *between* groups and to analysis of segments of social structure not easily defined in terms of primary groups.

A fundamental weakness of current sociological theory is that it does not relate micro-level interactions to macro-level patterns in any convincing way. Large-scale statistical, as well as qualitative, studies offer a good deal of insight into such macro phenomena as social mobility, community organization, and political structure. At the micro level, a large and increasing body of data and theory offers useful and illuminating ideas about what transpires within the confines of the small group. But how interaction in small groups aggregates to form large-scale patterns eludes us in most cases.

I will argue, in this paper, that the analysis of processes in interpersonal networks provides the most fruitful micro-macro bridge. In one way or another, it is through these networks that small-scale interaction becomes



Part 2

City indices

<https://unhabitat.org/wcr/>



What are good city indices?

(paste here examples)

Some past examples

<https://strelkamag.com/en/article/russian-cities-index>

Coding colab session

First rule: if you get errors and correct them, you are learning

Python typical errors:

ValueError: Nominatim could not geocode query to polygonal boundaries

Coding colab session

(looking into streets of Mumbai with python notebook)

<https://colab.research.google.com/drive/1Ey5Py9RiqNdJHjz643oOEiYhz5vPbyO4?usp=sharing>

```
[ ] # first let us convert MultiDiGraph into DiGraph
# Specify the name that is used to search for the data
import osmnx as ox
import networkx as nx

place_name = 'Mumbai, India' # Fetch OSM street network from the location
graphcity = ox.graph_from_place(place_name)# Get place boundary related to the place name as a geodataframe
area = ox.geocode_to_gdf(place_name)# Plot the streets
fig, ax = ox.plot_graph(graphcity)
# compose graph from several graphs in the area

#save G graph into file
print(type(graphcity))

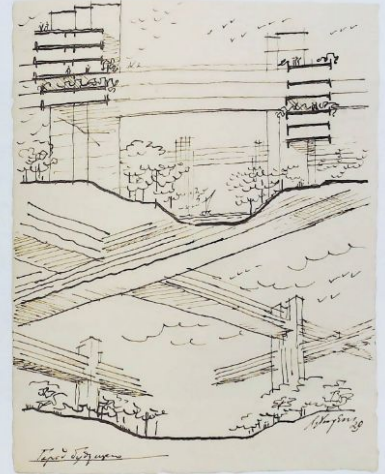
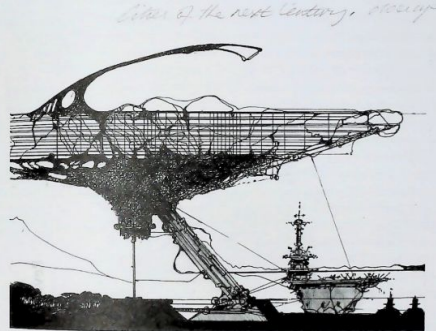
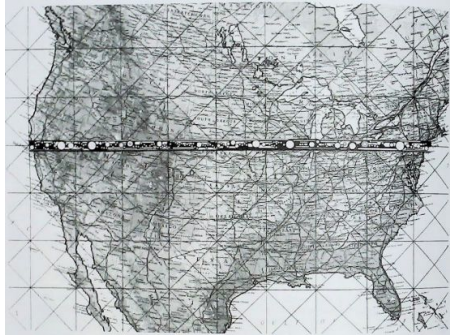
nx.write_edgelist(graphcity, "edgescity.csv", delimiter=" ") # edges with attributes!!!

edgestrastevere = graphcity.edges.data() # default data is {} (empty dict)

#print((edgestrastevere))
print(type(edgestrastevere))
#print(edgestrastevere[0])
```

Some books

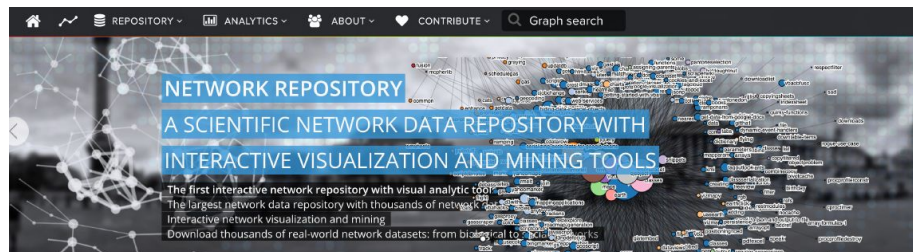
“The World as an Architectural Project”



Network resources

<http://networkrepository.com/networks.php>

<http://networksciencebook.com/chapter/3#advanced-b>



Network Repository. An Interactive *Scientific* Network Data Repository.

THE FIRST SCIENTIFIC NETWORK DATA REPOSITORY WITH INTERACTIVE VISUAL ANALYTICS.

NEW **GraphVis: interactive visual graph mining and machine learning**

The first interactive data and network data repository with real-time visual analytics. Network repository is not only the first interactive repository, but also the *largest network repository* with thousands of donations in 30+ domains (from biological to social network data). This large comprehensive collection of network graph data is useful for making significant research findings as well as benchmark network data sets for a wide variety of applications and domains (e.g., network science, bioinformatics, machine learning, data mining, physics, and social science) and includes relational, attributed, heterogeneous, streaming, spatial, and time series network data as well as non-relational machine learning data. All graph data sets are easily downloaded into a standard consistent format. We also have built a multi-level interactive graph analytics engine that allows users to visualize the structure of the network data as well as macro-level graph data statistics as well as important micro-level network properties of the nodes and edges.

Check out **GraphVis**: the interactive visual network mining and machine learning tool.

 GET NETWORK DATA

 COMPARE GRAPH DATA

 VISUALIZE NETWORKS

Network resources

<http://networkrepository.com/networks.php>

<http://networksciencebook.com/chapter/3#advanced-b>

Spatial Networks

Marc Barthélemy*

*Institut de Physique Théorique, CEA, IPhT CNRS, URA 2306 F-91191 Gif-sur-Yvette France and
Centre d'Analyse et de Mathématique Sociales (CAMS, UMR 8557 CNRS-EHESS)
Ecole des Hautes Etudes en Sciences Sociales, 54 bd. Raspail, F-75270 Paris Cedex 06, France.*

Complex systems are very often organized under the form of networks where nodes and edges are embedded in space. Transportation and mobility networks, Internet, mobile phone networks, power grids, social and contact networks, neural networks, are all examples where space is relevant and where topology alone does not contain all the information. Characterizing and understanding the structure and the evolution of spatial networks is thus crucial for many different fields ranging from urbanism to epidemiology. An important consequence of space on networks is that there is a cost associated to the length of edges which in turn has dramatic effects on the topological structure of these networks. We will expose thoroughly the current state of our understanding of how the spatial constraints affect the structure and properties of these networks. We will review the most recent empirical observations and the most important models of spatial networks. We will also discuss various processes which take place on these spatial networks, such as phase transitions, random walks, synchronization, navigation, resilience, and disease spread.

Contents

I. Networks and space	2	1. Erdos-Renyi graph	39
A. Introduction	2	2. Planar Erdos-Renyi graph	39
B. Quantitative geography and networks	2	3. The hidden variable model for spatial networks	40
C. What this review is (not) about	2	4. The Waxman model	41
II. Characterizing spatial networks	3	C. Spatial small worlds	41
A. Generalities on planar networks	3	1. The Watts-Strogatz model	41
1. Spatial and planar networks	3	2. Spatial generalizations	42
2. Classical results for planar networks	3	D. Spatial growth models	43
3. Voronoi tessellation	4	1. Generalities	43
		2. Preferential attachment and distance selection	43
		3. Growth and local optimization	46
		E. Optimal networks	49

