Network Science

2019. szeptember 9.

Technical infos
Contacts

Gergely Palla

e-mail:	**pallag@hal.elte.hu**
office:	Dept. of Biological Physics, Room 3.90. (North Block)
web:	http://hal.elte.hu/~pallag

Péter Pollner

e-mail:	**pollner@hal.elte.hu**
office:	Dept. of Biological Physics, Room 3.90. (North Block)
web:	http://hal.elte.hu/~pollner

Consultation time: Monday, 14-16.

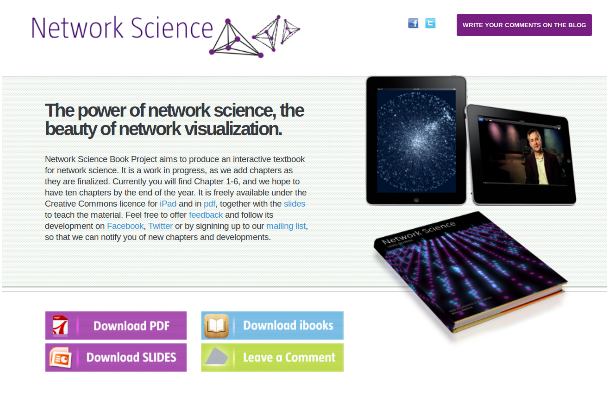
Online:

- Slides are always uploaded to the Moodle page of the course.

Technical infos
Reading material

Online:

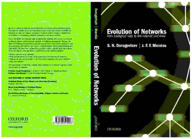
- Slides are always uploaded to the Moodle page of the course.
- Albert-László Barabási: Network Science
  http://networksciencebook.com



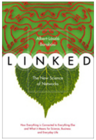
Technical infos
Reading material

Offline:

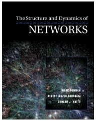
- J. F. F. Mendes and S. N. Dorogovtsev: **Evolution of Networks**, *(Oxford University Press, Oxford, 2003)*



- A.-L. Barabási: **Linked** *(Perseus Books Group, 2002)*



- M. E. J. Newman, A.-L. Barabási, D. J. Watts: **The Structure and Dynamics of Networks** *(Princeton University Press, Princeton, 2006)*



Grading

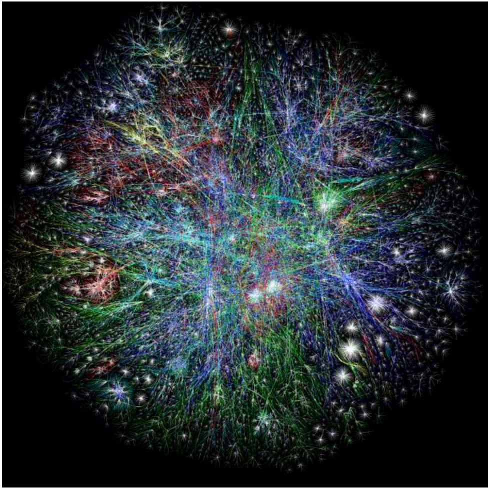
- Midterm test.
- Final test.
- Oral exam.

Introduction

BASIC QUESTIONS

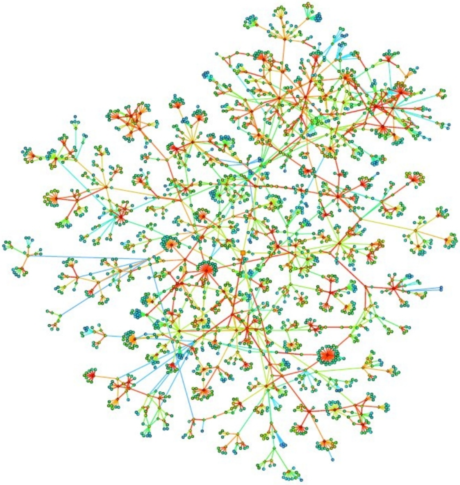
What are networks?

Networks...
(The Internet)

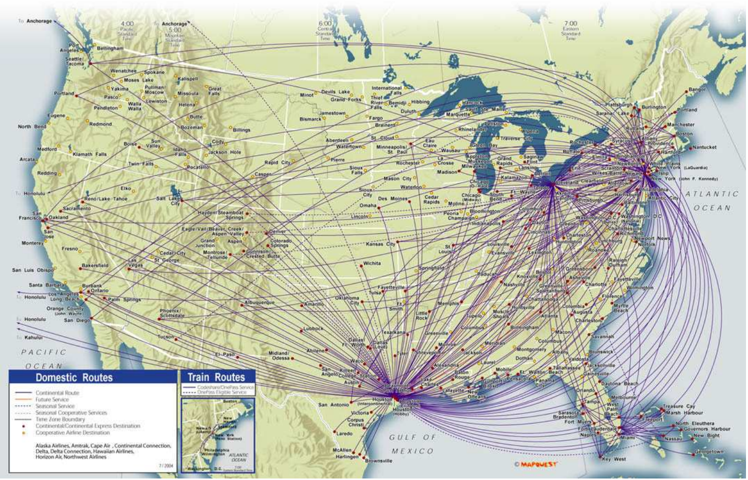


Networks...
(Social)

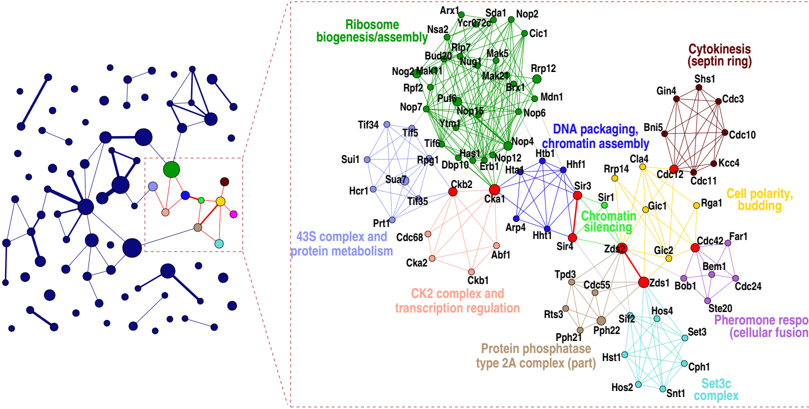
Networks...
(Social)



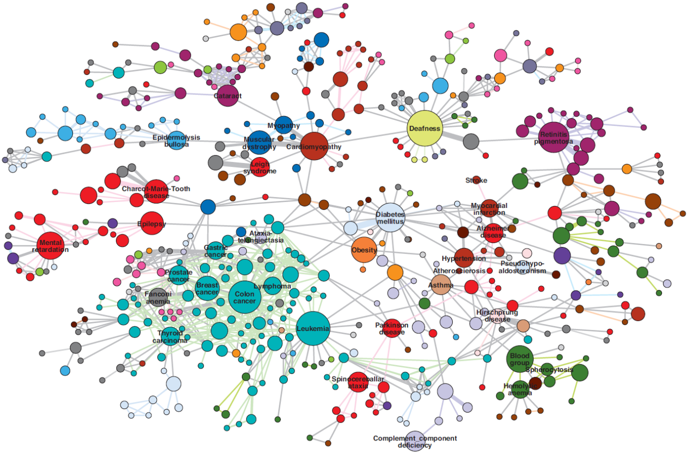
Networks...
(Transportation)



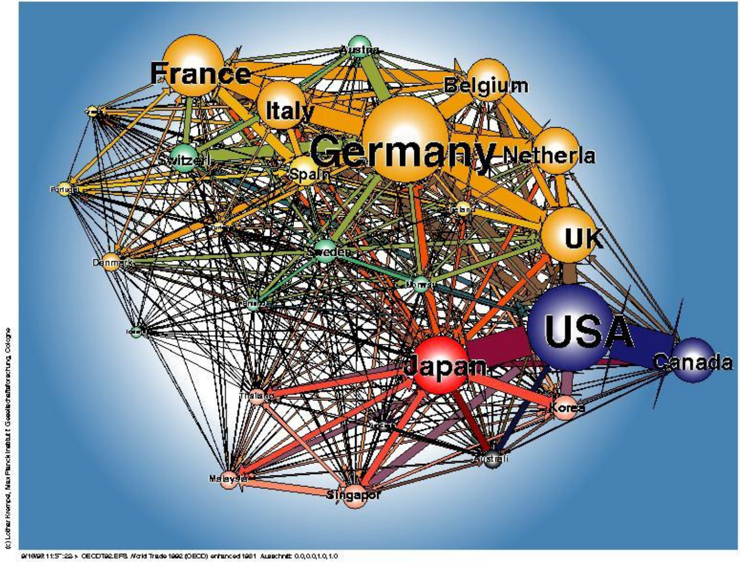
Networks...
(Protein interaction)



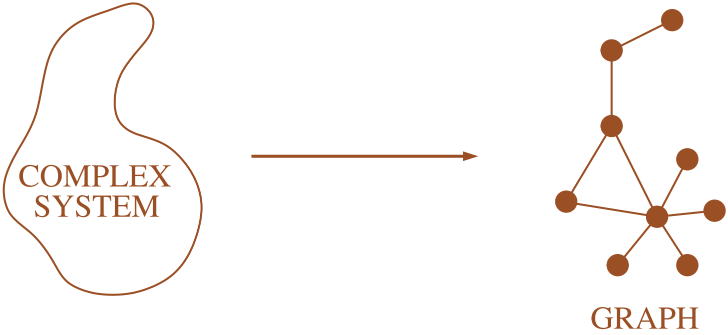
Networks...
(Human diseases)



Networks...
(World trade)



The network approach



What are complex systems?

Complex systems

Dictionary.com

com·plex [*adj., v.* kuhm-**pleks**, **kom**-pleks; *n.* **kom**-pleks] ?
Show IPA

adjective

1. composed of many interconnected parts; compound; composite: *a complex highway system.*
2. characterized by a very complicated or involved arrangement of parts, units, etc.: *complex machinery.*
3. so complicated or intricate as to be hard to understand or deal with: *a complex problem.*

WIKIPEDIA
The Free Encyclopedia

Complex system: Composed of interconnected parts that as a whole exhibit one or more properties (behavior among the possible properties) not obvious from the properties of the individual parts...

Complex networks

Are real networks complex?

- large size ✓
- intricate wiring ✓

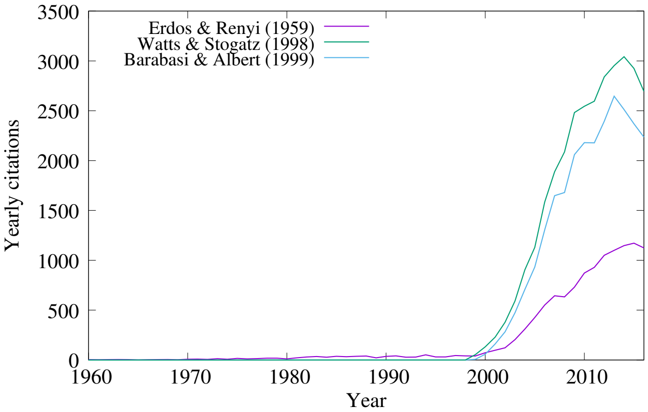
Why now?

Why now?

Citations of

- P. Erdős and A. Rényi, „On Random Graphs I", *Publ. Math. Debrecen* **6**, p. 290-297 (1959),
- D. J. Watts and S. H. Strogatz, „Collective dynamics of 'small-world' networks", *Nature* **393**, p. 440-442 (1998),

- A.-L. Barabási and R. Albert, „Emergence of scaling in random networks", *Science* **286**, p. 509-512 (1999)



Why is network theory one of the most dynamically evolving field of current times?

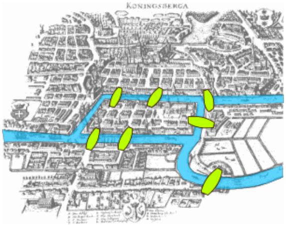
Why did we have to wait until the 2000s for the proliferation of network theory?

Why is network theory one of the most dynamically evolving field of current times?

Why did we have to wait until the 2000s for the proliferation of network theory?

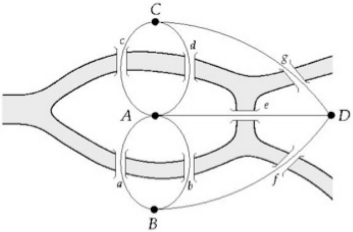
Historical overview of network research

- **Leonhard Euler** (1735):
  The bridges of Königsberg and graph theory.



Historical overview of network research

- **Leonhard Euler** (1735):
  The bridges of Königsberg and graph theory.



Historical overview of network research

- **Pál Erdős and Alfréd Rényi** (1959): Random graphs.

- **Stanley Milgram** (1967): Six degrees of separation.

Historical overview of network research

- **Duncan Watts and Steven H. Strogatz** (1998):
  Small world networks.


- **Albert-László Barabási and Réka Albert** (1999):
  Scale-free networks.


Why now?

Why now?

- **Data availability:**
  - Movie actor network 1998,
  - WWW 1999,
  - Citation network 1998,
  - Metabolic network 2000,
  - Protein interaction network 2001,
  - ...
- **Universality**
- **Applicability**

What are the applications of network theory?

Applications of network theory

Quantum Mechanics: 1900

- $\rightarrow$ electron microscope 1931,
- $\rightarrow$ transistor 1947,
- $\rightarrow$ laser 1957,
- $\rightarrow$ magnetic resonance imaging 1973,

In general, about a **30 year gap** between science and technology.

$\updownarrow$

In case of network science, this drops down to only about **3 years**!

Applications of network theory

Quantum Mechanics: 1900

- $\rightarrow$ electron microscope 1931,
- $\rightarrow$ transistor 1947,
- $\rightarrow$ laser 1957,
- $\rightarrow$ magnetic resonance imaging 1973,

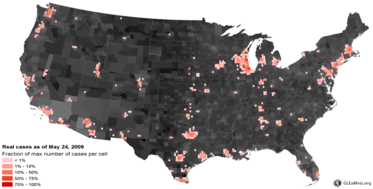
In general, about a **30 year gap** between science and technology.

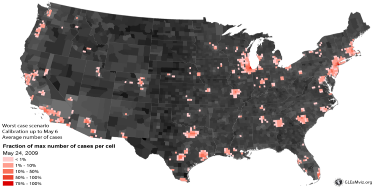
$\updownarrow$

In case of network science, this drops down to only about **3 years**!

Predicting the H1N1 pandemic in 2009:

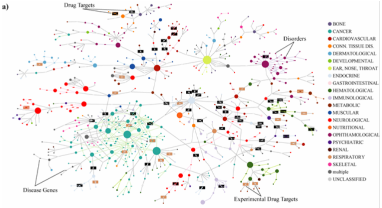
Real	**Projected**

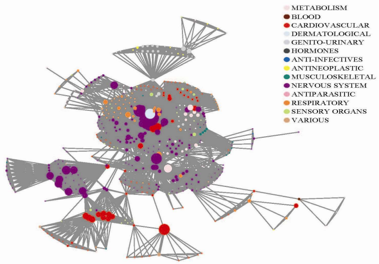


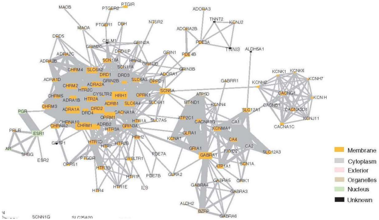


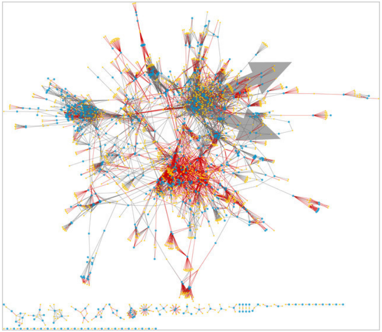
Applications of network theory

Drug design









Applications of network theory
Fighting terrorism

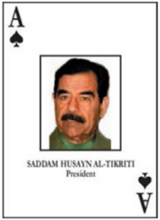
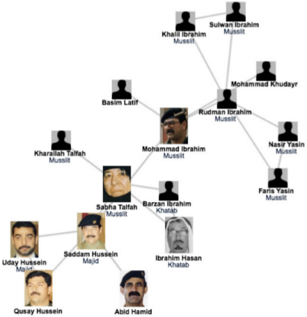


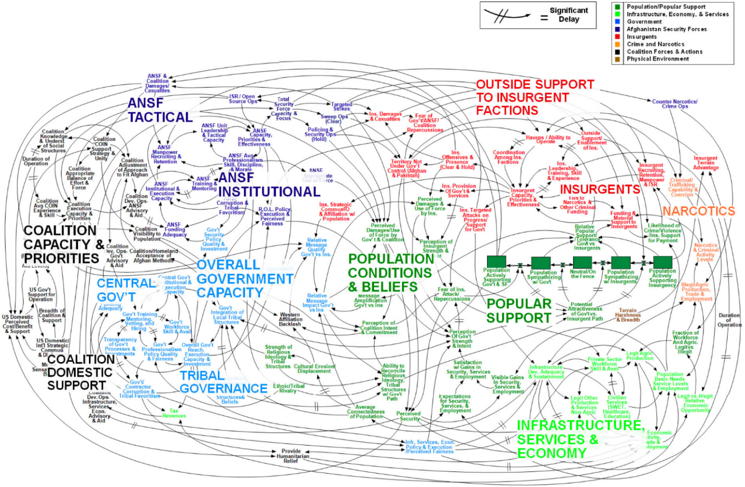
Image 1.2a
The network of Saddam Hussein.

Ace of Spades. One of the 55 cards the US military has handed out to the coalition forces in Iraq, each listing a top official to be captured following the country's 2003 invasion. The card shows the ace of spades, with the image of Saddam Hussein, Iraq's deposed president and dictator, the top prize of the hunt.



(From the book of A.-L. Barabási).

Applications of network theory
Fighting terrorism

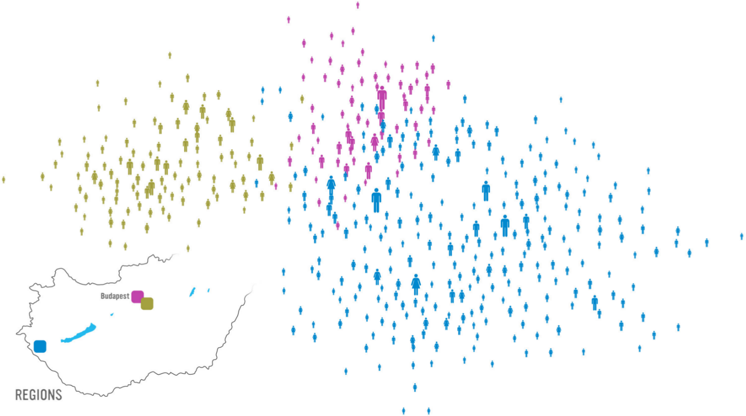


(From the book of A.-L. Barabási).

Applications of network theory
Fighting terrorism

http://www.ns-cta.org/ns-cta-blog/

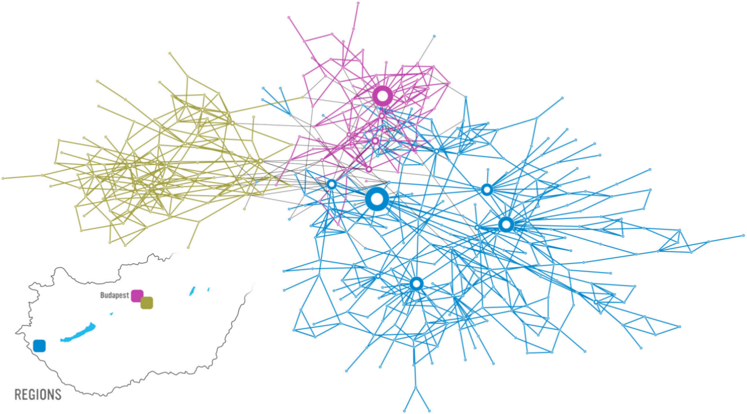
Applications of network theory
Organizational consulting



(From the book of A.-L. Barabási).

Applications of network theory

Organizational consulting



(From the book of A.-L. Barabási).

Applications of network theory
Organizational consulting

(From the book of A.-L. Barabási).

Applications of network theory
Organizational consulting

(From the book of A.-L. Barabási).

OVERVIEW

Outline

- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Outline

- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Outline

- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Basic network characteristics

- **How to characterize a node in a network?**

Basic network characteristics

- **How to characterize a node in a network?**
  - degree,
  - betweenness,
  - closeness,
  - etc.

Basic network characteristics

- **How to characterize a node in a network?**
  - degree,
  - betweenness,
  - closeness,
  - etc.

  $\rightarrow$ **Centralities!**

Basic network characteristics

- **How to characterize a node in a network?**
  - degree,
  - betweenness,
  - closeness,
  - etc.

  $\rightarrow$ **Centralities!**
- **Why are they important?**

Basic network characteristics

- **How to characterize a node in a network?**
  - degree,
  - betweenness,
  - closeness,
  - etc.

  → **Centralities!**
- **Why are they important?**

Outline

- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Outline

- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Advanced network characteristics

- **How to characterize a network?**

Advanced network characteristics

- **How to characterize a network?**
  - degree distribution,
  - clustering coefficient,
  - average shortest path length,
  - degree correlations,
  - etc.

Advanced network characteristics

- **How to characterize a network?**
  - degree distribution,
  - clustering coefficient,
  - average shortest path length,
  - degree correlations,
  - etc.

- **Why are they important?**

Advanced network characteristics

- **How to characterize a network?**
  - degree distribution,
  - clustering coefficient,
  - average shortest path length,
  - degree correlations,
  - etc.

- **Why are they important?**

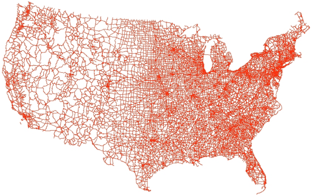
→ We can „categorize” networks.

Advanced network characteristics

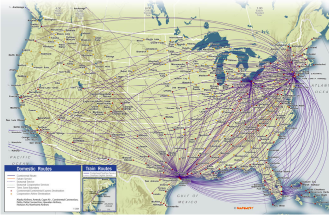
- **How to characterize a network?**
  - degree distribution,
  - clustering coefficient,
  - average shortest path length,
  - degree correlations,
  - etc.

- **Why are they important?**

→ We can „categorize” networks.



?



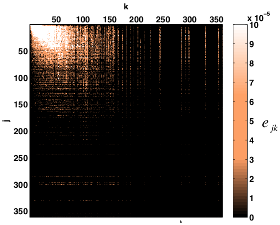
Advanced network characteristics

- **How to characterize a network?**
  - degree distribution,
  - clustering coefficient,
  - average shortest path length,
  - degree correlations,
  - etc.

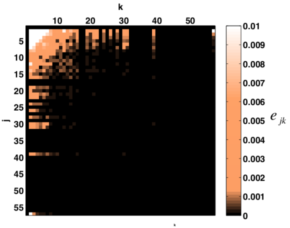
- **Why are they important?**

→ We can „categorize” networks.

Astrophys. co–authorship



Yeast PPI



Outline

- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Outline

- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Network models

- **How to generate networks?**
  - Erdős–Rényi model,
  - Watts-Strogatz model,
  - Barabási–Albert model,
  - Configuration model.
  - Hidden variable model.

- **Why are they important?**
  - Singling out essential features.
  - Explaining relevant phenomena via constructive mechanisms.
  - Generating test beds.

Outline

- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Outline

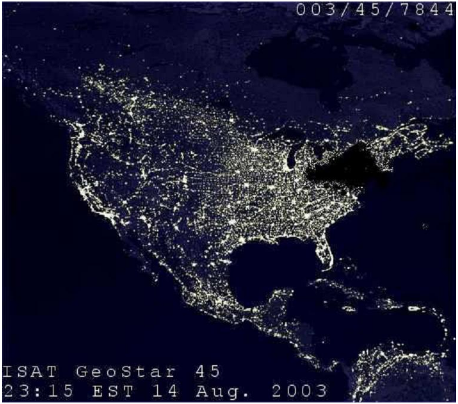
- Introduction
- Basic network characteristics
- Advanced network characteristics
- Network models
- Robustness and spreading
- Motifs and communities
- Hierarchical structure of complex networks

Robustness

- **Why are some networks more/less robust?**

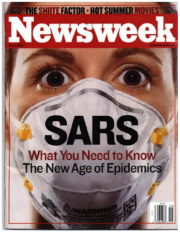
Robustness

- **Why are some networks more/less robust?**



  - How to describe in quantitative terms the breakdown of a network under node or link removal?
  - How does the network structure affect the robustness?

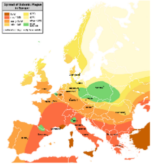
- **How do ideas, information or infections spread on networks?**



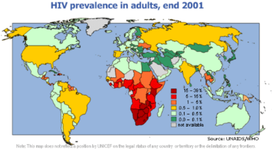
Spreading

- **How do ideas, information or infections spread on networks?**

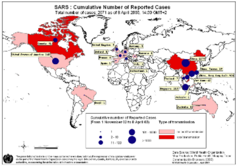
The Great Plague



HIV



SARS



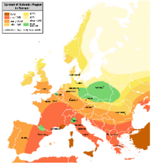
1918 Spanish flu

H1N1 flu

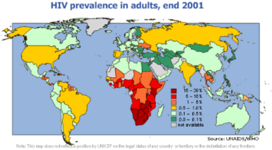
Spreading

- **How do ideas, information or infections spread on networks?**

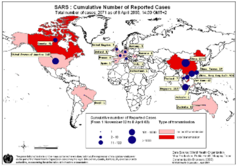
The Great Plague



HIV



SARS



1918 Spanish flu

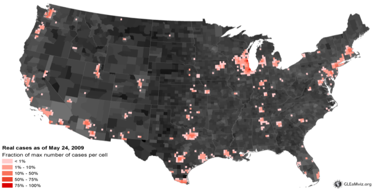
H1N1 flu

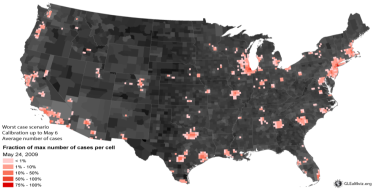
- How does the network structure affect the spreading rate?

Spreading

- **How do ideas, information or infections spread on networks?**

Real | **Projected**





- How does the network structure affect the spreading rate?

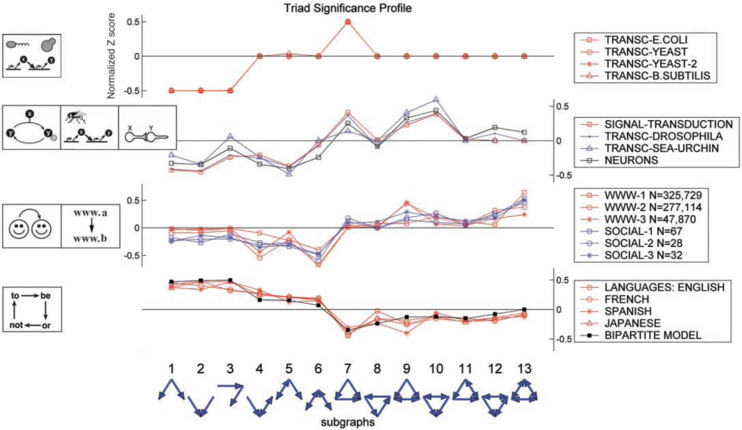
Outline

- Introduction
- Basic network characteristics
- Advanced network characteristics
- Network models
- Robustness and spreading
- Motifs and communities
- Hierarchical structure of complex networks

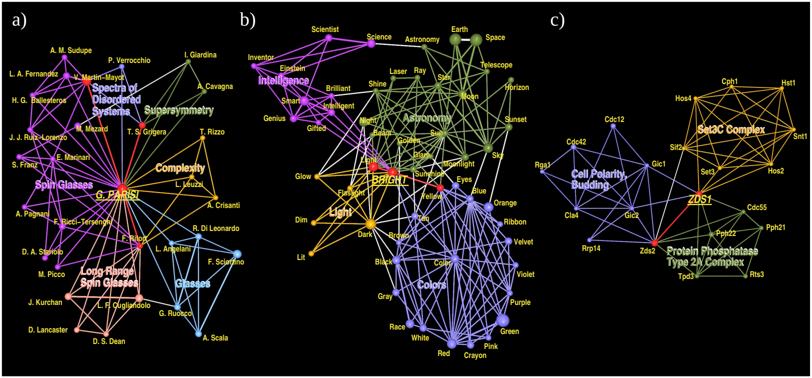
Outline

- Introduction
- Basic network characteristics
- Advanced network characteristics
- Network models
- Robustness and spreading
- Motifs and communities
- Hierarchical structure of complex networks

Motifs


Triad Significance Profile
Normalized Z score
0.5
0
-0.5
TRANSC-E.COLI
TRANSC-YEAST
TRANSC-YEAST-2
TRANSC-B.SUBTILIS
SIGNAL-TRANSDUCTION
TRANSC-DROSOPHILA
TRANSC-SEA-URCHIN
NEURONS
WWW-1 N=325,729
WWW-2 N=277,114
WWW-3 N=47,870
SOCIAL-1 N=67
SOCIAL-2 N=28
SOCIAL-3 N=32
LANGUAGES: ENGLISH
FRENCH
SPANISH
JAPANESE
BIPARTITE MODEL
www.a
www.b
to
be
not
or
1 2 3 4 5 6 7 8 9 10 11 12 13
subgraphs

Communities



Outline

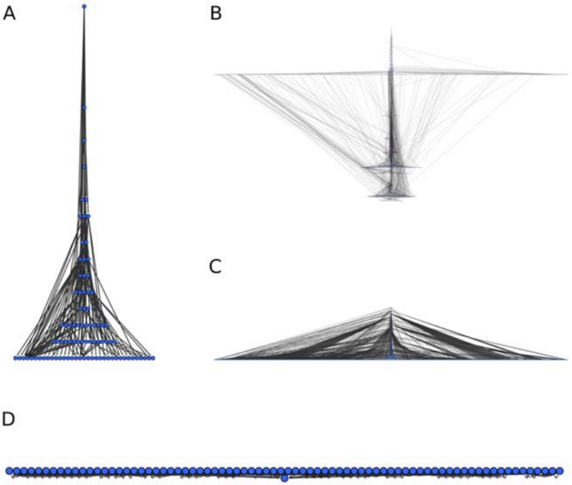
- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Outline

- **Introduction**
- **Basic network characteristics**
- **Advanced network characteristics**
- **Network models**
- **Robustness and spreading**
- **Motifs and communities**
- **Hierarchical structure of complex networks**

Hierarchical structure of complex networks

How to quantify the extent of hierarchy within the structure of a given network?



GRAPHS

Graph vs. network

Graph vs. network

Graph

- vertices,
- edges,
- possibly directed,
- possibly weighted

Network

- nodes,
- links,
- node props.: tags, annotations, state variable $x_i(t)$,
- flows,
- epidemics,
- growth,
- rewiring

Graph vs. network

Graph $\longleftrightarrow$ Network $\longleftrightarrow$ Complex system

Graph:

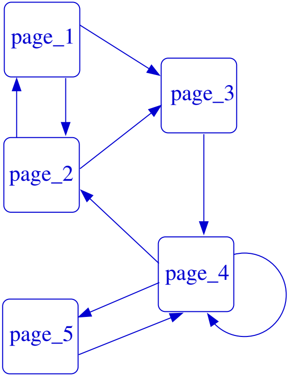
- vertices,
- edges,
- possibly directed,
- possibly weighted

Network:

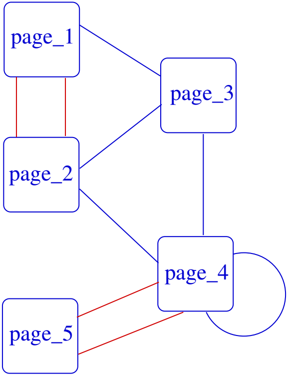
- nodes,
- links,
- node props.: tags, annotations, state variable $x_i(t)$,
- flows,
- epidemics,
- growth,
- rewiring

Kiddy games

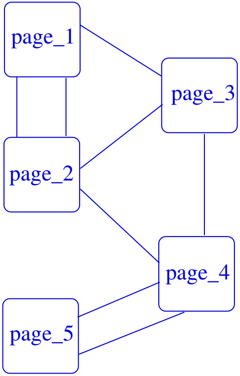
Representing the same system with different graphs



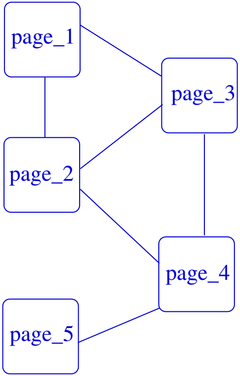
Representing the same system with different graphs



Representing the same system with different graphs



Representing the same system with different graphs



Graph types

Graph types

- simple graph: only single connections are allowed, no self loops.
- multi graph: multiple connections between the same pair of nodes are allowed, no self loops.
- pseudo graph: multiple connections and self loops are also allowed.

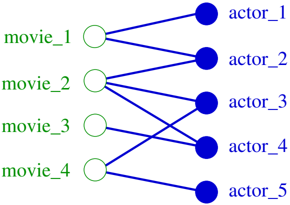
- directed graph: the links are directed.
- undirected graph: links are undirected.
- (mixed graph: contains both directed and undirected links).

- weighted graph: links have weight
- unweighted graph: links are „binary”

- bi-partite graph: two types of nodes, links are allowed only between opposite types.

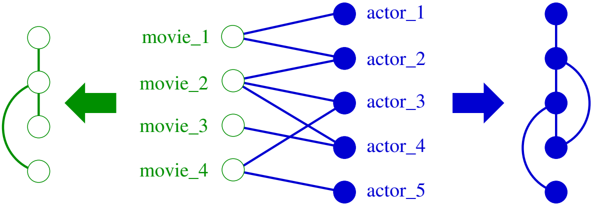
Bi-partite graphs

E.g., the network of movie actors:

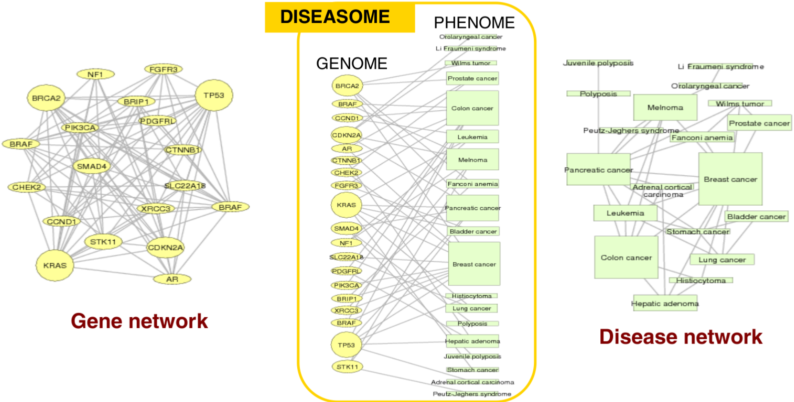


Bi-partite graphs

E.g., the network of movie actors:



Bi-partite graphs



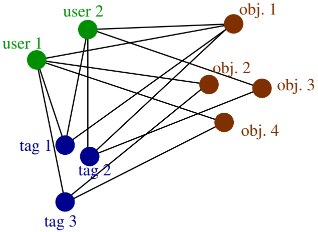
Goh, Cusick, Valle, Childs, Vidal & Barabási, PNAS (2007)

Tri-partite graphs?

Tri-partite graphs

Collaborative tagging systems (folksonomies):

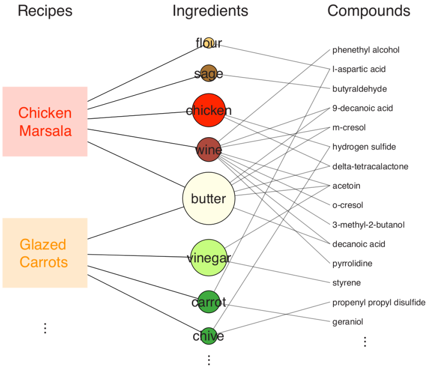
- A large group of users are attaching various tags to a collection of objects.
- Triplets of object-tag-user, $\longrightarrow$ tripartite graph.



- 6 different single-mode projections.

Tripartite graphs

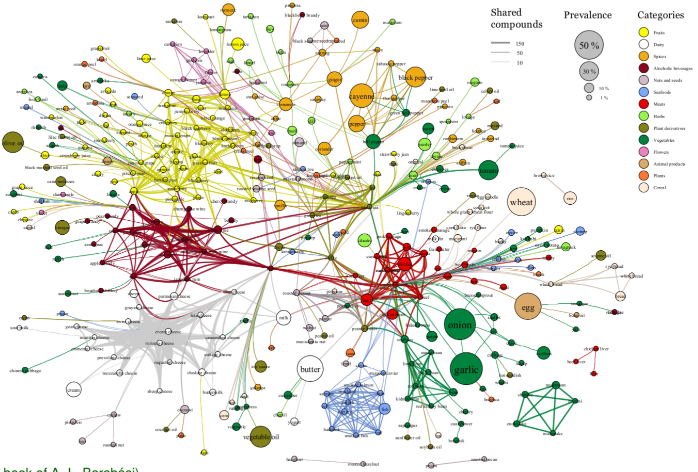
Recipe–ingredient–compound network:



(From the book of A.-L. Barabási)

Tripartite graphs

Recipe–ingredient–compound network:



(From the book of A.-L. Barabási)

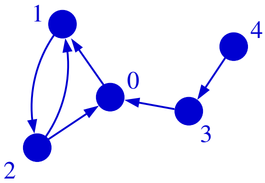
Network formats

How do we store a graph on the computer?

Adjacency matrix

- each row and column corresponds to a node,

$$A_{ij} = \begin{cases} 1 & \text{if } i \to j \\ 0 & \text{otherwise} \end{cases}$$



$\longleftrightarrow$?

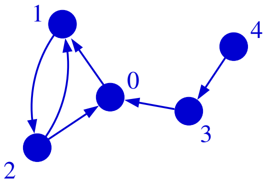
Adjacency matrix

Definition

Adjacency matrix

- each row and column corresponds to a node,

$$A_{ij} = \begin{cases} 1 & \text{if } i \to j \\ 0 & \text{otherwise} \end{cases}$$



$\longleftrightarrow$

$$\begin{array}{c c} & \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

Adjacency matrix
Properties

- if the graph is undirected,

Adjacency matrix

Properties

- if the graph is undirected, A_{ij} is symmetric.

Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs,

Adjacency matrix
Properties

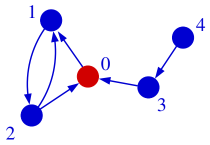
- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.

Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.
- Multiplying a position vector with A_{ij}:

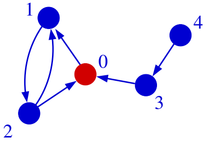
Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.
- Multiplying a position vector with A_{ij}:



Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.
- Multiplying a position vector with A_{ij}:

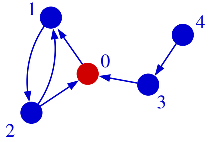


position:

$$\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}$$

Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.
- Multiplying a position vector with A_{ij}:

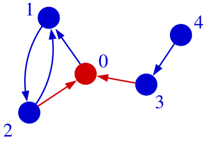


position:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = ?$$

Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.
- Multiplying a position vector with A_{ij}:

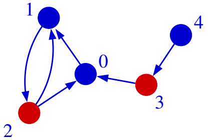


position:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.
- Multiplying a position vector with A_{ij}:

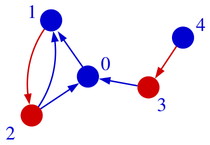


position:

$$\begin{pmatrix}0\\0\\1\\1\\0\end{pmatrix}$$

Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.
- Multiplying a position vector with A_{ij}:

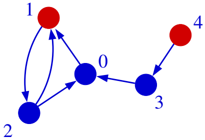


position:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.
- Multiplying a position vector with A_{ij}:

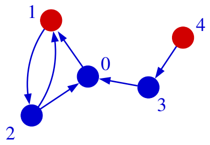


position:

$$\begin{pmatrix}0\\1\\0\\0\\1\end{pmatrix}$$

Adjacency matrix
Properties

- if the graph is undirected, A_{ij} is symmetric.
- for weighted graphs, A_{ij} can take real values.
- Multiplying a position vector with A_{ij}:



position:

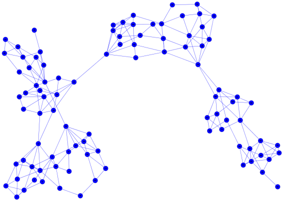
$$\begin{pmatrix}0\\1\\0\\0\\1\end{pmatrix}$$

- The spectrum of A_{ij} can be useful in various problems.

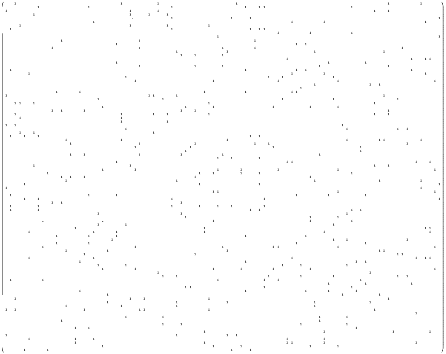
Adjacency matrix
Drawback

Adjacency matrix
Drawback

- The A_{ij} for real networks contains mostly zeros...

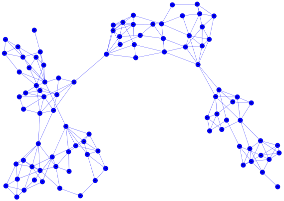


$\leftrightarrow$

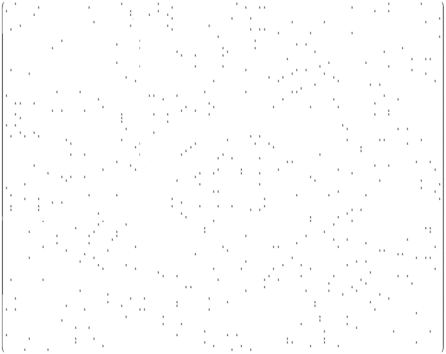


Adjacency matrix
Drawback

- The A_{ij} for real networks contains mostly zeros...



$\leftrightarrow$



- This is called sparseness.

Adjacency matrix
Drawback

- Storing a lot of zeros does not seem to be a good idea...

Adjacency matrix
Drawback

- Storing a lot of zeros does not seem to be a good idea...

→ What to do?

Adjacency matrix
Drawback

- Storing a lot of zeros does not seem to be a good idea...

→ What to do?

- Use a link list instead.

Sparseness

How would you give sparseness a mathematical definition?

Sparseness

Sparseness

- A network (graph) is **sparse** in the $N \to \infty$ limit (N denotes the number of nodes) **if the number of links** $M \sim N$.
- A network (graph) is **dense** in the $N \to \infty$ limit (N denotes the number of nodes) **if the number of links** $M \sim N^2$.

What about the average number of connections per node?

Sparseness

Sparseness

- A network (graph) is **sparse** in the $N \to \infty$ limit (N denotes the number of nodes) **if the number of links** $M \sim N$.
- A network (graph) is **dense** in the $N \to \infty$ limit (N denotes the number of nodes) **if the number of links** $M \sim N^2$.

What about the average number of connections per node?

Sparseness

Sparseness

- A network (graph) is **sparse** in the $N \to \infty$ limit (N denotes the number of nodes) **if the number of links** $M \sim N$.
- A network (graph) is **dense** in the $N \to \infty$ limit (N denotes the number of nodes) **if the number of links** $M \sim N^2$.

What about the average number of connections per node?

- sparse case: $\langle k \rangle = \frac{2M}{N} \to \mathrm{const.}$,
- dense case: $\langle k \rangle = \frac{2M}{N} \sim N \to \infty$!

What does this mean in practice?

(How would you decide whether a real network of finite size is sparse or not?)

Sparseness
In practice

What does this mean in practice?

(How would you decide whether a real network of finite size is sparse or not?)

If $\langle k \rangle << N$ **by several orders of magnitude smaller**, then the network is considered **sparse**.

Sparseness

Most real networks are sparse!

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference	Nr.
WWW, site level, undir.	153, 127	35.21	3.1	3.35	0.1078	0.00023	Adamic 1999	1
Internet, domain level	3015 - 6209	3.52 - 4.11	3.7 - 3.76	6.36 - 6.18	0.18 - 0.3	0.001	Yook *et al.* 2001a, Pastor-Satorras *et al.* 2001	2
Movie actors	225, 226	61	3.65	2.99	0.79	0.00027	Watts, Strogatz 1998	3
LANL coauthorship	52, 909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman 2001a,b	4
MEDLINE coauthorship	1, 520, 251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman 2001a,b	5
SPIRES coauthorship	56, 627	173	4.0	2.12	0.726	0.003	Newman 2001a,b,c	6
NCSTRL coauthorship	11, 994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman 2001a,b	7
Math coauthorship	70, 975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.* 2001	8
Neurosci. coauthorship	209, 293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási *et al.* 2001	9
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner, Fell 2000	10
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner, Fell 2000	11
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya, Solé 2000	12
Silwood park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya, Solé 2000	13
Words, cooccurence	460.902	70.13	2.67	3.03	0.437	0.0001	Cancho, Solé 2001	14
Words, synonyms	22, 311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.* 2001	15
Power grid	4, 941	2.67	18.7	12.4	0.08	0.005	Watts, Strogatz 1998	16
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts, Strogatz 1998	17

What are the consequences of sparseness?

- we use a **link list instead of the adjacency matrix**,
- $\langle k \rangle << N$

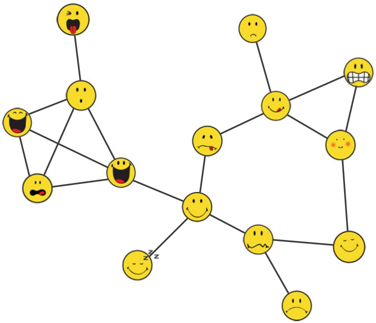
What are the consequences of sparseness?

- we use a **link list instead of the adjacency matrix**,
- $\langle k \rangle << N$
- the probability for a connection between a randomly chosen pair?

What are the consequences of sparseness?

- we use a **link list instead of the adjacency matrix**,
- $\langle k \rangle << N$
- for a randomly chosen pair of nodes, **the probability of being linked is negligible!** (I.e., the probability converges to 0 if $N \to \infty$).

Basic network characteristics

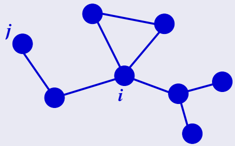


BASIC NODE PROPERTIES

Node degree

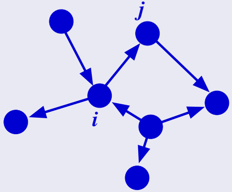
Node degree

- Number of connections, denoted by k_i or d_i.



$$k_i = 4,\ k_j = 1.$$

- Directed networks $\longrightarrow$ k_i^{in} and k_i^{out}.

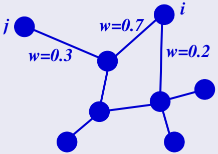


$$k_i^{\text{in}} = 2,\ k_j^{\text{out}} = 1.$$

Node degree

Node degree

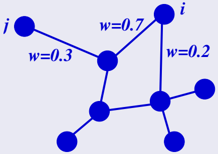
- Weighted networks?



Node degree

Node degree

- Weighted networks: node strength, s_i



$$s_i = 0.9,\ s_j = 0.3.$$

Node degree

Examples

What is the degree of node i in the following examples?

Node degree
Examples

What is the degree of node i in the following examples?

- a 1d chain of nodes? 

Node degree
Examples

What is the degree of node i in the following examples?

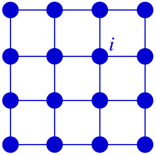
- a 1d chain of nodes? 

$k_i = 2$

Node degree
Examples

What is the degree of node i in the following examples?

- a 1d chain of nodes?  $k_i = 2$
- a 2d square lattice?

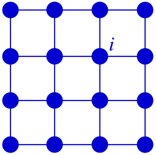


Node degree

Examples

What is the degree of node i in the following examples?

- a 1d chain of nodes?  $k_i = 2$
- a 2d square lattice?



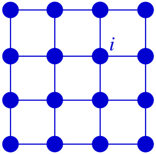
$k_i = 4$

Node degree
Examples

What is the degree of node i in the following examples?

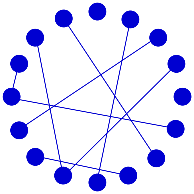
- a 1d chain of nodes? 

$k_i = 2$
- a 2d square lattice?



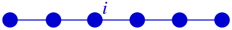
$k_i = 4$

- a random graph of N nodes, where the uniform linking probability is p ? (Erdős-Rényi model)

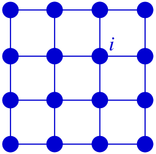


Node degree
Examples

What is the degree of node i in the following examples?

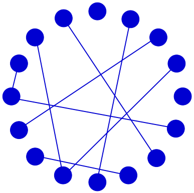
- a 1d chain of nodes? 

$k_i = 2$
- a 2d square lattice?



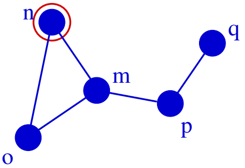
$k_i = 4$

- a random graph of N nodes, where the uniform linking probability is p ? (Erdős-Rényi model)



$\langle k_i \rangle = (N-1)p \simeq Np$

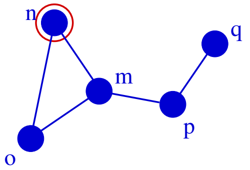
Node degree and adjacency matrix



$$\longleftrightarrow \quad \begin{array}{c} \\ m \\ n \\ o \\ p \\ q \end{array} \begin{array}{c} \begin{array}{ccccc} m & n & o & p & q \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

$k_{\mathrm{n}} \quad = \quad ?$

Node degree and adjacency matrix

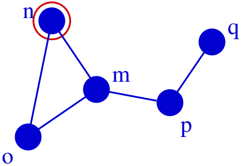


$$\longleftrightarrow$$

$$\begin{array}{c} \\ m \\ n \\ o \\ p \\ q \end{array} \begin{array}{c} \begin{array}{ccccc} m & n & o & p & q \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

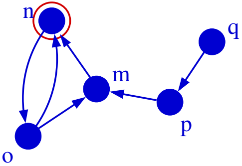
$$\begin{aligned} k_{\mathrm{n}} &= \sum_j A_{jn} \\ &= \sum_j A_{nj} \end{aligned}$$

Node degree and adjacency matrix



$$
\begin{array}{c|ccccc}
 & m & n & o & p & q \\
\hline
m & 0 & 1 & 1 & 1 & 0 \\
n & 1 & 0 & 1 & 0 & 0 \\
o & 1 & 1 & 0 & 0 & 0 \\
p & 1 & 0 & 0 & 0 & 1 \\
q & 0 & 0 & 0 & 1 & 0
\end{array}
$$

$$
\begin{aligned}
k_{\mathrm{n}} &= \sum_j A_{jn} \\
&= \sum_j A_{nj}
\end{aligned}
$$

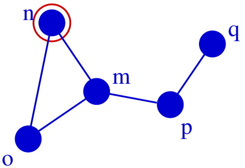


$$
\begin{array}{c|ccccc}
 & m & n & o & p & q \\
\hline
m & 0 & 1 & 0 & 0 & 0 \\
n & 0 & 0 & 1 & 0 & 0 \\
o & 1 & 1 & 0 & 0 & 0 \\
p & 1 & 0 & 0 & 0 & 0 \\
q & 0 & 0 & 0 & 1 & 0
\end{array}
$$

$$k_{\mathrm{n}}^{\mathrm{in}} = ?$$

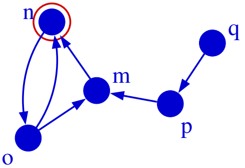
$$k_{\mathrm{n}}^{\mathrm{out}} = ?$$

Node degree and adjacency matrix



$$\longleftrightarrow \quad \begin{array}{c} \\ m \\ n \\ o \\ p \\ q \end{array} \begin{array}{c} \begin{array}{ccccc} m & n & o & p & q \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

$$k_{\mathrm{n}} = \sum_j A_{jn} = \sum_j A_{nj}$$



$$\longleftrightarrow \quad \begin{array}{c} \\ m \\ n \\ o \\ p \\ q \end{array} \begin{array}{c} \begin{array}{ccccc} m & n & o & p & q \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

$$k_{\mathrm{n}}^{\mathrm{in}} = \sum_j A_{jn}$$

$$k_{\mathrm{n}}^{\mathrm{out}} = \sum_j A_{nj}$$

Average degree

Average degree

Average degree

$$\langle k \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} k_i = \frac{2M}{N},$$

- a sort of density.
- directed networks:

$$\langle k_{\text{in}} \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} k_{i,\text{in}} = \frac{M}{N} = \langle k_{\text{out}} \rangle$$

Triangles, transitivity, clustering

- Is the friend of your friend a friend of yours as well?
- Do two of your friends know each other as well? (Are they friends?)

Triangles, transitivity, clustering

- Is the friend of your friend a friend of yours as well?
- Do two of your friends know each other as well? (Are they friends?)

  In many cases yes.

Triangles, transitivity, clustering

- Is the friend of your friend a friend of yours as well?
- Do two of your friends know each other as well? (Are they friends?)

  In many cases yes.

$\rightarrow$ lot of **triangles** in the network of acquaintances.

Triangles, transitivity, clustering

- Is the friend of your friend a friend of yours as well?
- Do two of your friends know each other as well? (Are they friends?)

  In many cases yes.

→ lot of **triangles** in the network of acquaintances.

Quantifying local transitivity: **clustering coefficient**.

Clustering coefficient

Clustering coefficient

- The clustering coefficient of node i:

$$C_i \equiv \frac{2e_i}{k_i(k_i-1)},$$

where e_i stands for the number of links between its neighbors. (Thus, $C_i \in [0,1]$).

- If $k_i < 2$, then $C_i = 0$ by definition.

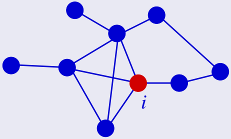
Clustering coefficient

Clustering coefficient

- The clustering coefficient of node i:

$$C_i \equiv \frac{2e_i}{k_i(k_i-1)},$$

where e_i stands for the number of links between its neighbors.
(Thus, $C_i \in [0,1]$).
Example:



- If $k_i < 2$, then $C_i = 0$ by definition.

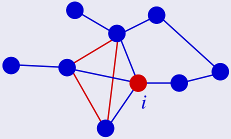
Clustering coefficient

Clustering coefficient

- The clustering coefficient of node i:

$$C_i \equiv \frac{2e_i}{k_i(k_i-1)},$$

where e_i stands for the number of links between its neighbors. (Thus, $C_i \in [0,1]$).
Example:



$$C_i = \frac{2e_i}{k_i(k_i-1)} = \frac{2\cdot 3}{4\cdot 3} = \frac{2}{4} = \frac{1}{2}$$

- If $k_i < 2$, then $C_i = 0$ by definition.

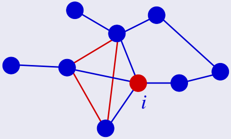
Clustering coefficient

Clustering coefficient

- The clustering coefficient of node i:

$$C_i \equiv \frac{2e_i}{k_i(k_i-1)},$$

where e_i stands for the number of links between its neighbors. (Thus, $C_i \in [0,1]$).
Example:



$$C_i = \frac{2e_i}{k_i(k_i-1)} = \frac{2\cdot 3}{4\cdot 3} = \frac{2}{4} = \frac{1}{2}$$

- If $k_i < 2$, then $C_i = 0$ by definition.

Clustering coefficient
Directed networks

Clustering coefficient in directed networks

- In a directed network twice as many links are possible between the neighbors of a given node, thus,

$$C_i \equiv \frac{e_i}{k_i(k_i - 1)}.$$

Average clustering and transitivity coefficient

Average clustering coefficient

- The average clustering coefficient of the network is

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i,$$

where N denotes the number of nodes in the network.

Clustering coefficient
Illustration

- If $\langle k \rangle$ is a measure of the **global** density of a network, what does $\langle C \rangle$ measure?

Clustering coefficient

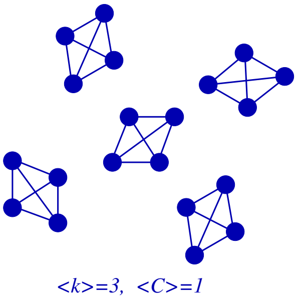
Illustration

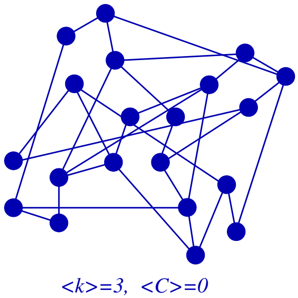
- If $\langle k \rangle$ is a measure of the **global** density of a network, what does $\langle C \rangle$ measure?

  $\longrightarrow$ **Local** density!

- If $\langle k \rangle$ is a measure of the **global** density of a network, what does $\langle C \rangle$ measure?

  $\longrightarrow$ **Local** density!





C in real networks

The $\langle C \rangle$ in real networks and their randomized counterparts (with the same N and M):

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference
WWW, site level, undir.	153, 127	35.21	3.1	3.35	0.1078	0.00023	Adamic 1999
Internet, domain level	3015 - 6209	3.52 - 4.11	3.7 - 3.76	6.36 - 6.18	0.18 - 0.3	0.001	Yook *et al.* 2001a, Pastor-Satorras *et al.* 2001
Movie actors	225, 226	61	3.65	2.99	0.79	0.00027	Watts, Strogatz 1998
LANL coauthorship	52, 909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman 2001a,b
MEDLINE coauthorship	1, 520, 251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman 2001a,b
SPIRES coauthorship	56, 627	173	4.0	2.12	0.726	0.003	Newman 2001a,b,c
NCSTRL coauthorship	11, 994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman 2001a,b
Math coauthorship	70, 975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.* 2001
Neurosci. coauthorship	209, 293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási *et al.* 2001
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner, Fell 2000
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner, Fell 2000
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya, Solé 2000
Silwood park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya, Solé 2000
Words, cooccurence	460.902	70.13	2.67	3.03	0.437	0.0001	Cancho, Solé 2001
Words, synonyms	22, 311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.* 2001
Power grid	4, 941	2.67	18.7	12.4	0.08	0.005	Watts, Strogatz 1998
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts, Strogatz 1998

C in real networks

The $\langle C \rangle$ in real networks and their randomized counterparts (with the same N and M):

Network	Size	$\langle k \rangle$	ℓ	ℓ_{rand}	C	C_{rand}	Reference
WWW, site level, undir.	153,127	35.21	3.1	3.35	0.1078	0.00023	Adamic 1999
Internet, domain level	3015 - 6209	3.52 - 4.11	3.7 - 3.76	6.36 - 6.18	0.18 - 0.3	0.001	Yook *et al.* 2001a, Pastor-Satorras *et al.* 2001
Movie actors	225,226	61	3.65	2.99	0.79	0.00027	Watts, Strogatz 1998
LANL coauthorship	52,909	9.7	5.9	4.79	0.43	1.8×10^{-4}	Newman 2001a,b
MEDLINE coauthorship	1,520,251	18.1	4.6	4.91	0.066	1.1×10^{-5}	Newman 2001a,b
SPIRES coauthorship	56,627	173	4.0	2.12	0.726	0.003	Newman 2001a,b,c
NCSTRL coauthorship	11,994	3.59	9.7	7.34	0.496	3×10^{-4}	Newman 2001a,b
Math coauthorship	70,975	3.9	9.5	8.2	0.59	5.4×10^{-5}	Barabási *et al.* 2001
Neurosci. coauthorship	209,293	11.5	6	5.01	0.76	5.5×10^{-5}	Barabási *et al.* 2001
E. coli, substrate graph	282	7.35	2.9	3.04	0.32	0.026	Wagner, Fell 2000
E. coli, reaction graph	315	28.3	2.62	1.98	0.59	0.09	Wagner, Fell 2000
Ythan estuary food web	134	8.7	2.43	2.26	0.22	0.06	Montoya, Solé 2000
Silwood park food web	154	4.75	3.40	3.23	0.15	0.03	Montoya, Solé 2000
Words, cooccurence	460.902	70.13	2.67	3.03	0.437	0.0001	Cancho, Solé 2001
Words, synonyms	22,311	13.48	4.5	3.84	0.7	0.0006	Yook *et al.* 2001
Power grid	4,941	2.67	18.7	12.4	0.08	0.005	Watts, Strogatz 1998
C. Elegans	282	14	2.65	2.25	0.28	0.05	Watts, Strogatz 1998

MOST REAL NETWORKS ARE HIGHLY CLUSTERED!

Why is C so much lower in the random networks?

Why is C so much lower in the random networks?

- Interpretation of C: probability that a pair of your neighbors are linked to each other.

Why is C so much lower in the random networks?

- Interpretation of C: probability that a pair of your neighbors are linked to each other.

$\rightarrow$ If the network is **sparse** and **random**, this probability is very small...

DISTANCE AND PATHS

Distance

Distance

We need to define a path first...

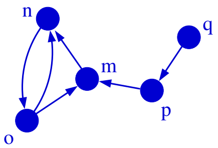
Pathology

Path

- A sequence of nodes in which each node is adjacent to the next one.

$\rightarrow$ it may intersect itself,

$\rightarrow$ in directed networks it must follow the direction of the links.



E.g., a legitimate path:
$p \rightarrow m \rightarrow n \rightarrow o \rightarrow m \rightarrow n$

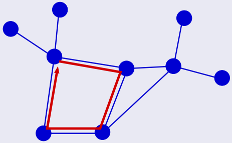
Pathology
Special paths

Pathology

Special paths

Cycle

A path with the same start and end node



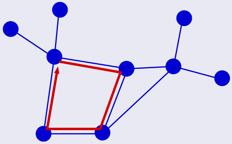
Self avoiding path

A path that does not intersect itself.

Pathology
Special paths

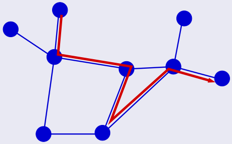
Cycle

A path with the same start and end node



Self avoiding path

A path that does not intersect itself.



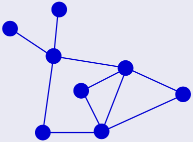
Pathology
Exotic paths

Pathology

Exotic paths

Eulerian path

A path traversing each link exactly once.



(E.g., the problem of the bridges of Königsberg.)

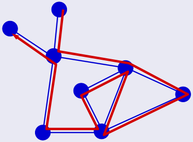
Hamiltonian path

A path visiting each node exactly once.

Pathology
Exotic paths

Eulerian path

A path traversing each link exactly once.



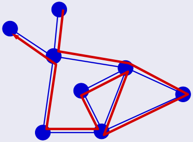
(E.g., the problem of the bridges of Königsberg.)

Hamiltonian path

A path visiting each node exactly once.

Eulerian path

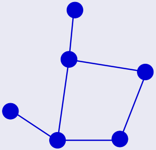
A path traversing each link exactly once.



(E.g., the problem of the bridges of Königsberg.)

Hamiltonian path

A path visiting each node exactly once.

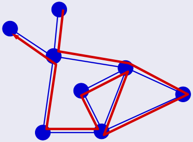


Pathology

Exotic paths

Eulerian path

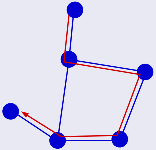
A path traversing each link exactly once.



(E.g., the problem of the bridges of Königsberg.)

Hamiltonian path

A path visiting each node exactly once.



Shortest path

Distance (shortest path length)

- ℓ_{ij}: minimum number of steps it takes to reach j from i following the links.
- Properties:
  - $\ell_{ii} = 0$
  - In undirected networks $\ell_{ij} = \ell_{ji}$.
  - If we cannot reach j from i, then $\ell_{ij} = \infty$.
- weighted networks?

Shortest path

Distance (shortest path length)

- ℓ_{ij}: minimum number of steps it takes to reach j from i following the links.
- Properties:
  - $\ell_{ii} = 0$
  - In undirected networks $\ell_{ij} = \ell_{ji}$.
  - If we cannot reach j from i, then $\ell_{ij} = \infty$.
- weighted networks?

Shortest path

Distance (shortest path length)

- ℓ_{ij}: minimum number of steps it takes to reach j from i following the links.
- Properties:
  - $\ell_{ii} = 0$
  - In undirected networks $\ell_{ij} = \ell_{ji}$.
  - If we cannot reach j from i, then $\ell_{ij} = \infty$.
- weighted networks: minimum weight path.

Shortest path

Average distance (average shortest path length)

- For a given node i:

$$\langle \ell_i \rangle = \frac{1}{N-1} \sum_j \ell_{ij},$$

for the whole network

$$\langle \ell \rangle = \frac{2}{N(N-1)} \sum_{i<j} \ell_{ij}$$

- $\langle \ell \rangle$ is also called as the average shortest path length between nodes.

Shortest path

Average distance (average shortest path length)

- For a given node i:

$$\langle \ell_i \rangle = \frac{1}{N-1} \sum_j \ell_{ij},$$

  for the whole network

$$\langle \ell \rangle = \frac{2}{N(N-1)} \sum_{i<j} \ell_{ij}$$

- $\langle \ell \rangle$ is also called as the average shortest path length between nodes.

Calculating shortest path length

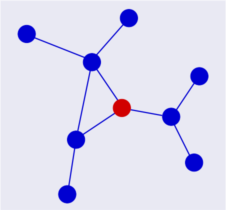
Calculating shortest path length

- breadth first search,
- depth first search,
- Dijkstra’s algorithm

Breadth first search

Breadth first search

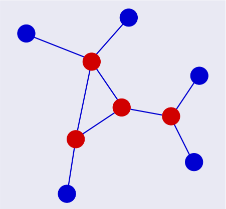
- Start from the root node, explore all of its neighbors.
- Next, explore all unvisited neighbors of the first neighbors.
- And so on, continue until no more unvisited nodes can be reached.



Breadth first search

Breadth first search

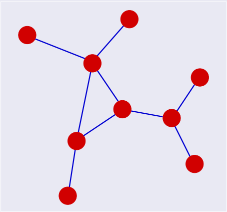
- Start from the root node, explore all of its neighbors.
- Next, explore all unvisited neighbors of the first neighbors.
- And so on, continue until no more unvisited nodes can be reached.



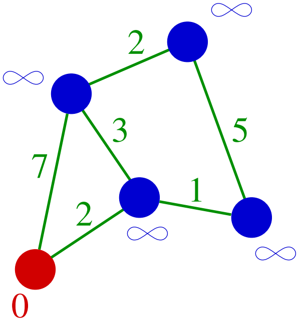
Breadth first search

Breadth first search

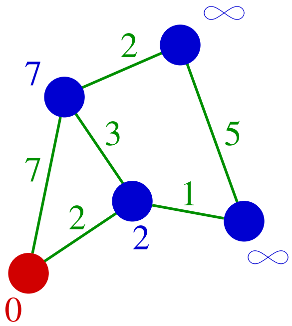
- Start from the root node, explore all of its neighbors.
- Next, explore all unvisited neighbors of the first neighbors.
- And so on, continue until no more unvisited nodes can be reached.



Dijkstra’s algorithm
Illustration

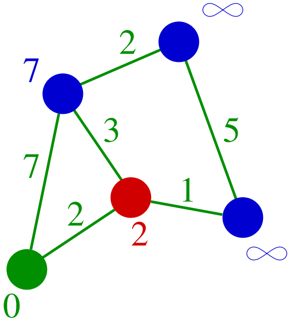


Dijkstra’s algorithm
Illustration



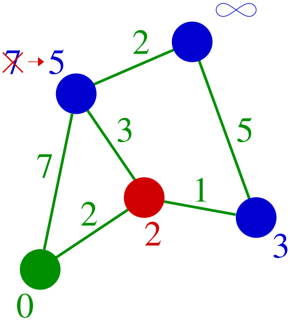
Dijkstra's algorithm

Illustration

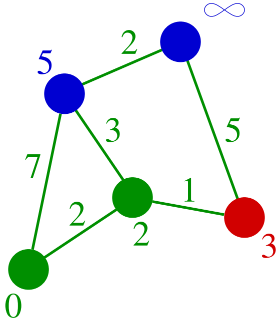


Dijkstra's algorithm

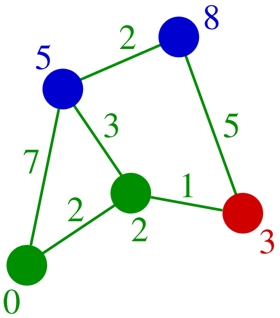
Illustration



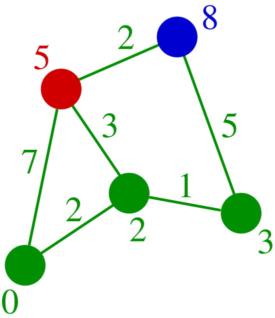
Dijkstra's algorithm
Illustration



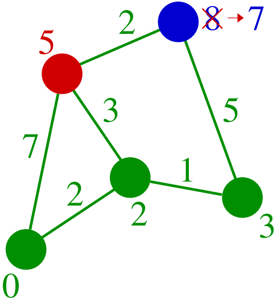
Dijkstra's algorithm
Illustration



Dijkstra's algorithm
Illustration

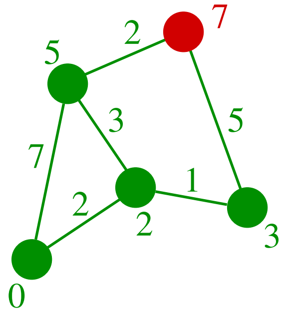


Dijkstra’s algorithm
Illustration

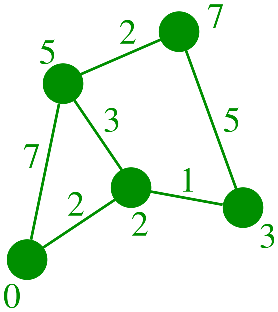


Dijkstra’s algorithm

Illustration



Dijkstra's algorithm
Illustration



Dijkstra's algorithm

Dijkstra's algorithm

This algorithm calculates the minimum weight paths to all nodes from a given source node:

- Assign to every node a weight: set it to zero for the source node and to infinity for all other nodes.
- Mark all nodes as unvisited. Set source node as current.
- For current node, calculate the tentative weight of its unvisited neighbors by adding the link weight to the weight of the current node. If this weight is less than the previously recorded weight on the neighbor, (infinity in the beginning), overwrite the weight.
- When done with considering all neighbors of the current node, mark it as visited. A visited node will not be checked ever again; its weight is already minimal.
- Set the unvisited node with the smallest weight to be the next current node. If no more unvisited nodes are left, finish.
- The complexity of the algorithm for sparse graphs can be estimated to be $\mathcal{O}(M \log(N))$.

Diameter

Diameter

Diameter

- The diameter of a network is given by the maximum distance between any pair of nodes,

$$D = \max_{ij} \ell_{ij}.$$

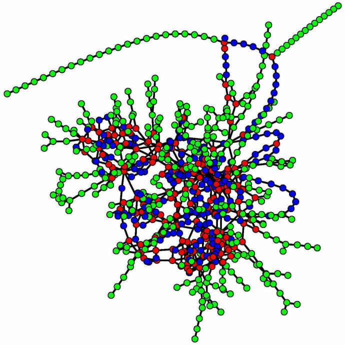
Diameter
Drawback

Basic network characteristics

Node properties
- Degree
- Clustering

Distance and paths
- Pathology
- Distance
- The small world property

Usually the average shortest path length is more descriptive:



The world is small...

Network of acquaintances:

The world is small...

Network of acquaintances:

- Occasionally we are surprised by the fact that people we think to have a large distance between each other are actually directly linked,

The world is small...

Network of acquaintances:

- Occasionally we are surprised by the fact that people we think to have a large distance between each other are actually directly linked,
- and $\langle \ell \rangle$ between people is $\mathcal{O}(10)$, whereas the network size is $\mathcal{O}(10^7)$.

The world is small...

Network of acquaintances:

- Occasionally we are surprised by the fact that people we think to have a large distance between each other are actually directly linked,
- and $\langle \ell \rangle$ between people is $\mathcal{O}(10)$, whereas the network size is $\mathcal{O}(10^7)$.
- How do we know?

The world is small...

Network of acquaintances:

- Occasionally we are surprised by the fact that people we think to have a large distance between each other are actually directly linked,
- and $\langle \ell \rangle$ between people is $\mathcal{O}(10)$, whereas the network size is $\mathcal{O}(10^7)$.
- How do we know?

$\longrightarrow$ The Milgram experiment.

Frigyes Karinthy

The six degrees of separation was first suggested by **Frigyes Karinthy**, a Hungarian writer in the 1920s...

(1887-1938)

Chains (1929)
"The worker knows the manager in the shop, who knows Ford; Ford is on friendly terms with the general director of Hearst Publications, who last year became good friends with Árpád Pásztor, someone I not only know, but to the best of my knowledge a good friend of mine. So I could easily ask him to send a telegram via the general director telling Ford that he should talk to the manager and have the worker in the shop quickly hammer together a car for me, as I happen to need one."

The Milgram experiment (1967)

HOW TO TAKE PART IN THIS STUDY

1. ADD YOUR NAME TO THE ROSTER AT THE BOTTOM OF THIS SHEET, so that the next person who receives this letter will know who it came from.
2. DETACH ONE POSTCARD. FILL IT AND RETURN IT TO HARVARD UNIVERSITY. No stamp is needed. The postcard is very important. It allows us to keep track of the progress of the folder as it moves toward the target person.
3. IF YOU KNOW THE TARGET PERSON ON A PERSONAL BASIS, MAIL THIS FOLDER DIRECTLY TO HIM (HER). Do this only if you have previously met the target person and know each other on a first name basis.
4. IF YOU DO NOT KNOW THE TARGET PERSON ON A PERSONAL BASIS, DO NOT TRY TO CONTACT HIM DIRECTLY. INSTEAD, MAIL THIS FOLDER (POST CARDS AND ALL) TO A PERSONAL ACQUAINTANCE WHO IS MORE LIKELY THAN YOU TO KNOW THE TARGET PERSON. You may send the folder to a friend, relative or acquaintance, but it must be someone you know on a first name basis.

The Milgram experiment

Results of the experiment:

- Surprisingly a part of the letters did reach the target, and the average number of steps needed was around 6.
- Thus, a rough estimate of $\langle l \rangle$ for people in the USA is around 6, which is several orders of magnitude smaller than the size of the population...
- Indeed, the world is small.