Random networks and small worlds

Pierre Senellart (pierre.senellart@telecom-paristech.fr)

Athens course on *Collective Intelligence*
18 March 2009

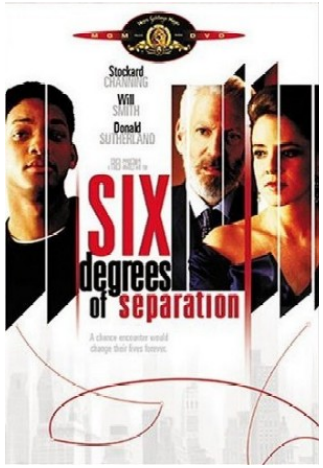
I proposed a more difficult problem: to find a chain of contacts linking myself with an anonymous riveter at the Ford Motor Company — and I accomplished it in four steps. The worker knows his foreman, who knows Mr. Ford himself, who, in turn, is on good terms with the director general of the Hearst publishing empire. I had a close friend, Mr. Árpád Pásztor, who had recently struck up an acquaintance with the director of Hearst Publishing. It would take but one word to my friend to send a cable to the general director of Hearst asking him to contact Ford who could in turn contact the foreman, who could then contact the riveter, who could then assemble a new automobile for me, would I need one.

[...] Our friend was absolutely correct: nobody from the group needed more than five links in the chain to reach, just by using the method of acquaintance, any inhabitant of our Planet.

[Karinthy, 1929]

Six Degrees of Separation

- Idea that two persons on Earth are separated by a chain of six individuals who know each other
- Appears widely in popular culture:



It's a small world!

Stanley Milgram's Experiment [Travers and Milgram, 1969]

- Stanley Milgram (1933-1984): social psychologist
- Experiment: people are asked to send a message to some unknown person, by forwarding it to an acquaintance who might be closer to this person
- Results: only 29% of the messages arrived, with a mean number of acquaintances of 5.2.
- Validates somehow the 6-degree theory!
- Other more recent experiments [Dodds et al., 2003] confirm this order of magnitude.

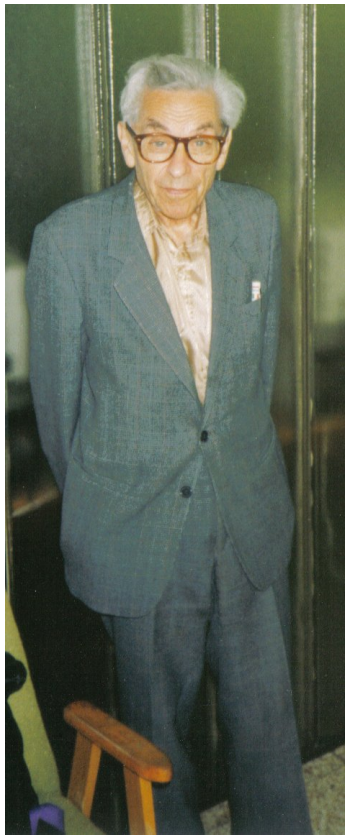
Kevin Bacon's Number

(David Shankbone, Wikimedia)

- Kevin Bacon: Hollywood actor, played in numerous movies, mostly secondary roles
- Kevin Bacon's number:
  - 0 for Kevin Bacon himself
  - 1 for actors who played in the same movie as Bacon
  - 2 for actors who played in the same movie as someone with a number of 1
  - etc.
- `http://oracleofbacon.org/`

Most actors have a small Bacon's number!

Erdős number



(Kmhkmh, Wikimedia)

- Paul Erdős (1913-1996):
  Mathematician and computer scientist, worked across many fields, with may collaborators
- Erdős number:
  - 0 for Paul Erdős himself
  - 1 for scientists who coauthored an article with Erdős
  - 2 for scientists who coauthored an article with someone with a number of 1
  - etc.
- `http://www.ams.org/mathscinet/collaborationDistance.html`

Most scientists have a small Erdős number!

Questions

- Is there really a pattern here?
- How can this be mathematically modeled?
- Can we explain what happens?
- Anything else to discover in such networks?

Outline

1 Introduction

2 Basics of Graph Theory

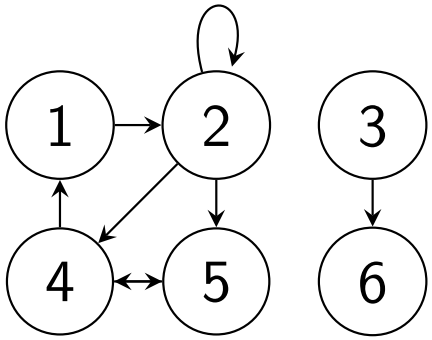
3 Characteristics of Real-World Networks

4 Models of Networks

5 Graph Mining Algorithms

6 Conclusion

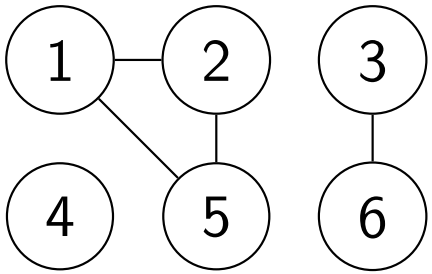
Graphs



Definition

A directed graph is a pair (S, A) where:

- S is a finite set of vertices (or nodes)
- A is a subset of S^2 defining the edges (or arcs)



Definition

An undirected graph is a pair (S, A) where:

- S is a finite set of vertices (or nodes)
- A is a set of (unordered) pairs of elements of S defining the edges (or arcs)

Remark

Graph is the mathematical term, network is used to describe real-world graphs.

Paths and Connectedness

Definition

A path is a sequence of vertices $v_1 \dots v_n$ such that v_k is connected by an edge to v_{k+1} for $1 \leq k \leq n-1$.

Definition

The underlying undirected graph of a directed graph G is the graph obtained by adding all reverse edges.

Definition

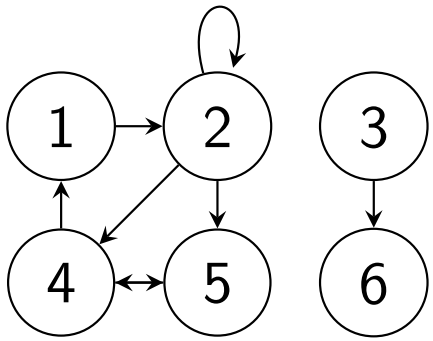
An undirected graph is connected if for every two vertices u and v, there exists a path starting from u and ending in v.
A directed graph is strongly connected if it is connected, and is weakly connected if the underlying undirected graph is connected.

Connected Components

Definition

(S', A') is a subgraph of (S, A) if $S' \subseteq S$ and A' is the restriction of A to edges whose vertices are in S'.

- Connected component: maximal connected subgraph
- Strongly connected component: maximal strongly connected subgraph
- Weakly connected component: maximal weakly connected subgraph

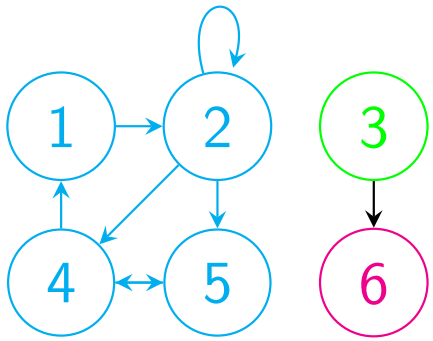


Connected Components

Definition

(S', A') is a subgraph of (S, A) if $S' \subseteq S$ and A' is the restriction of A to edges whose vertices are in S'.

- Connected component: maximal connected subgraph
- Strongly connected component: maximal strongly connected subgraph
- Weakly connected component: maximal weakly connected subgraph



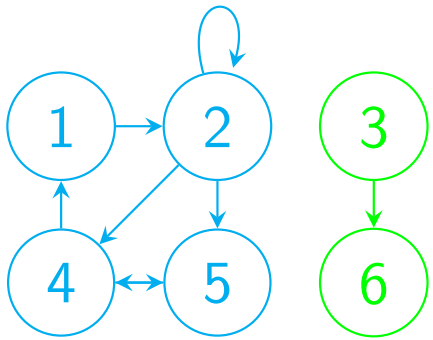
Strongly connected components

Connected Components

Definition

(S', A') is a subgraph of (S, A) if $S' \subseteq S$ and A' is the restriction of A to edges whose vertices are in S'.

- Connected component: maximal connected subgraph
- Strongly connected component: maximal strongly connected subgraph
- Weakly connected component: maximal weakly connected subgraph



Weakly connected components

Vocabulary

Incident: an edge is said to be incident to a vertex if it it has this vertex for endpoint

Degree (of a vertex): number of edges incident to a vertex, in an undirected graph

Indegree (of a vertex): number of edges arriving to a vertex, in a directed graph

Outdegree (of a vertex): number of edges leaving from a vertex, in a directed graph

Cycle: Path whose start and end vertex is the same

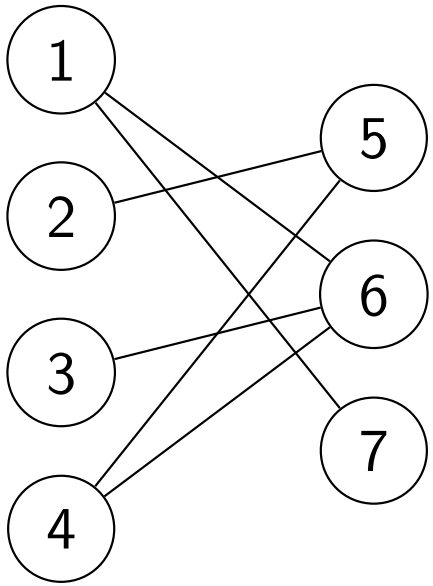
Distance: Length of the shortest path between two vertices

Sparse: a graph (S, A) is sparse if $|A| \ll |S|^2$

Bipartite Graphs

Definition

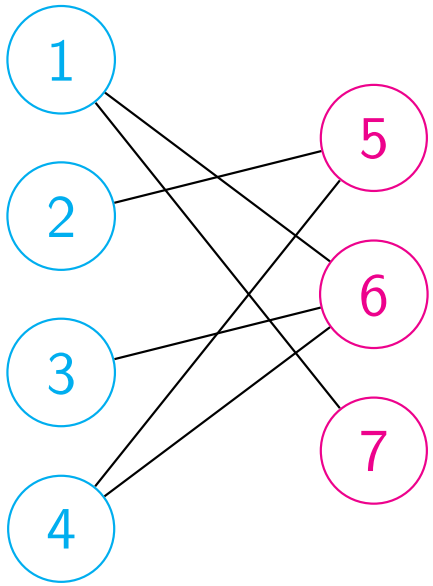
A bipartite graph is an undirected graph (S, A) such that $S = S_1 \cup S_2$ (with $S_1 \cap S_2 = \varnothing$), and no edge of A is incident to two vertices in S_1 or two vertices in S_2.



Bipartite Graphs

Definition

A bipartite graph is an undirected graph (S, A) such that $S = S_1 \cup S_2$ (with $S_1 \cap S_2 = \varnothing$), and no edge of A is incident to two vertices in S_1 or two vertices in S_2.

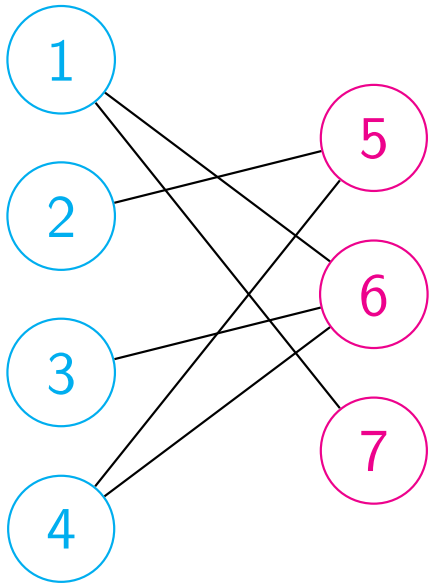


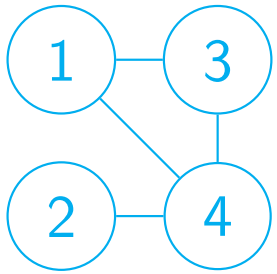
Bipartite Graphs

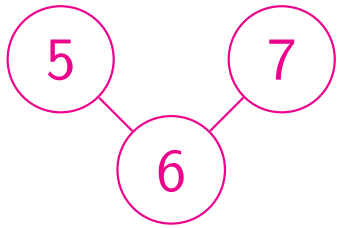
Definition

A bipartite graph is an undirected graph (S, A) such that $S = S_1 \cup S_2$ (with $S_1 \cap S_2 = \varnothing$), and no edge of A is incident to two vertices in S_1 or two vertices in S_2.

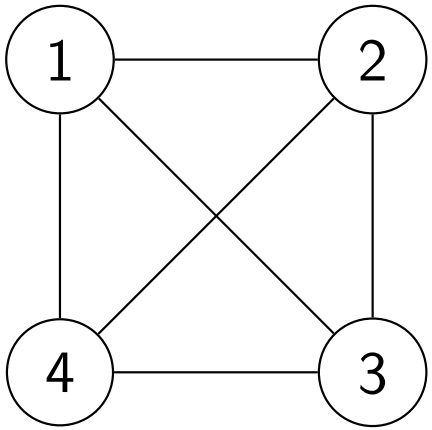
Paths of length 2 in a bipartite graph define two regular undirected graphs.

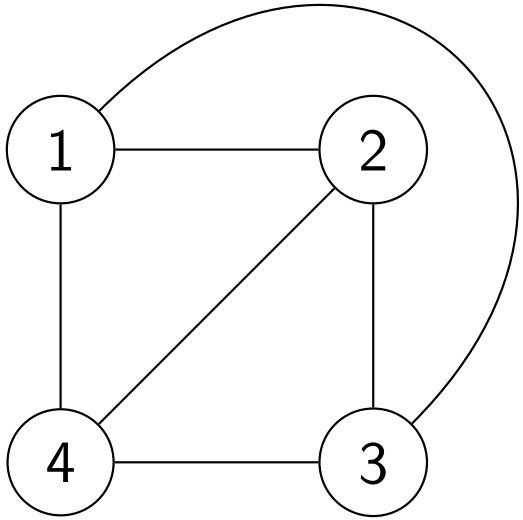


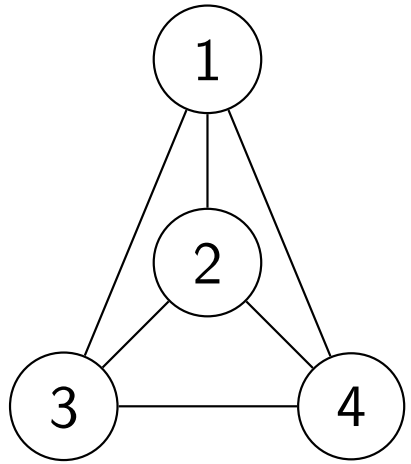




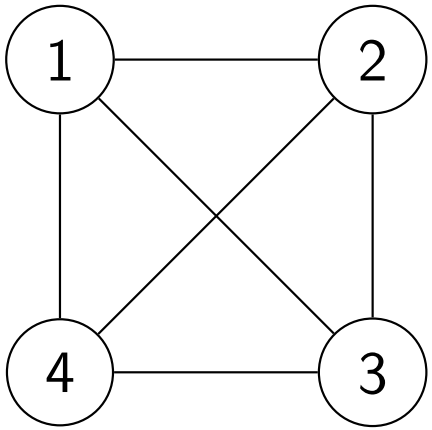
Beware of Graph Drawings

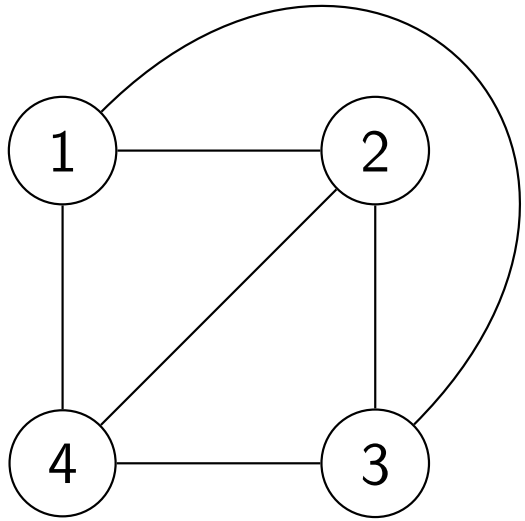


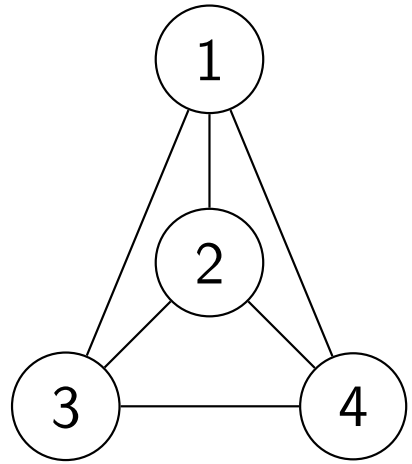




Beware of Graph Drawings





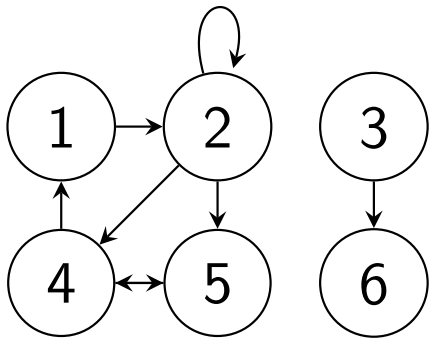


Three times the same graph!

- No "best" graph
- Not always possible to have a planar graph

Matrix Representation of a Graph

A graph G can be represented by its adjacency matrix M:



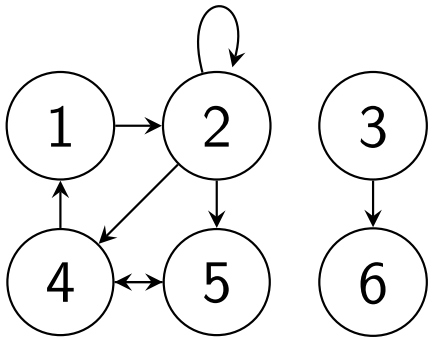
$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- Adjacency matrices of undirected graphs are symmetric.
- M^{T} is the adjacency matrix obtained from G by reversing the arrows.
- M^n is the matrix of the graph of all paths of length n in G.

Concrete Representation of Graphs

In programming, graphs are usually represented as:

- its adjacency matrix (stored as a multidimensional array), for non-sparse graphs
- the list of all edges incident to each node, for sparse graphs



1	2
2	2, 4, 5
3	6
4	1, 5
5	4
6	

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks
- Social Networks
- Natural Networks
- Artificial Networks

4 Models of Networks

5 Graph Mining Algorithms

6 Conclusion

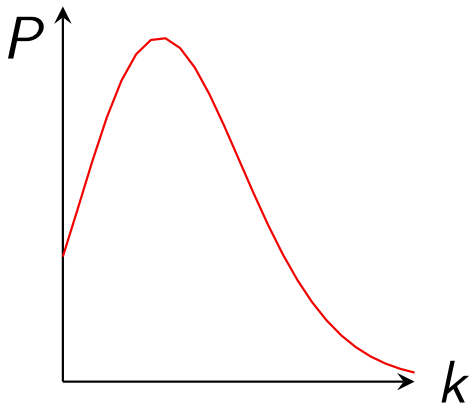
Characteristics of Interest

Sparsity. Is the network sparse ($|A| \ll |S|^2$)?
All networks considered here will be sparse.

Typical distance. What is the mean distance between any pairs of vertices?

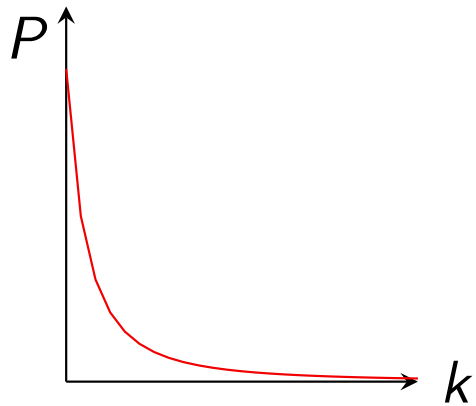
Local clustering. If a is connected to both b and c, is the probability that b is connected to c significantly greater than the probability any two nodes are connected?

Degree distribution. What is the distribution of the degree of vertices?



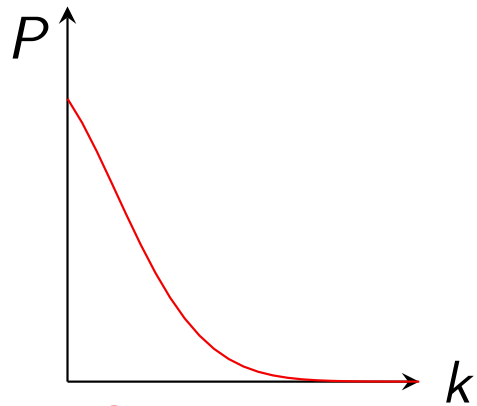
Poisson

$$\frac{\lambda^k}{k!}$$



Power-law

$$k^{-\gamma}$$



Gaussian

$$e^{-k^2}$$

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks
- Social Networks
- Natural Networks
- Artificial Networks

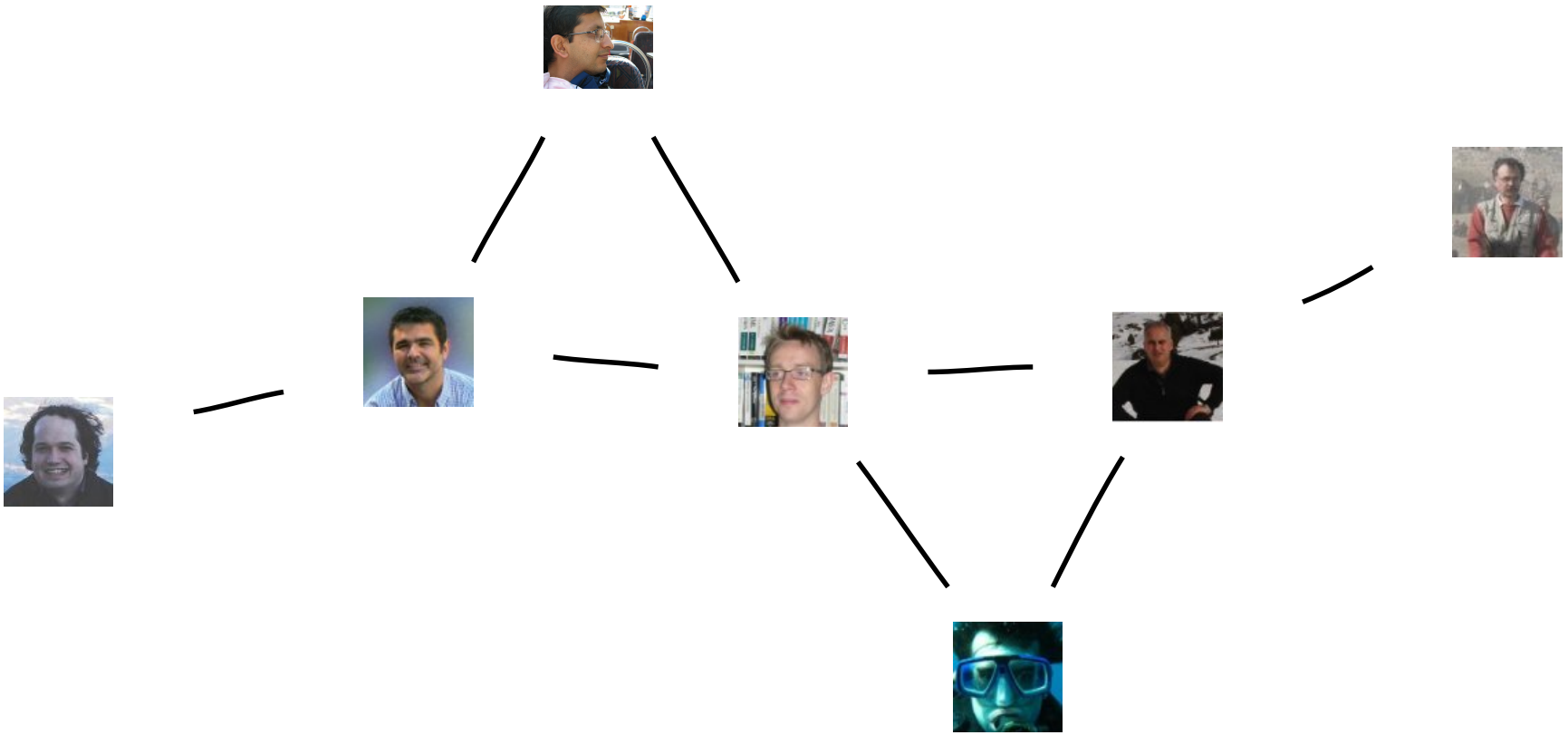
4 Models of Networks

5 Graph Mining Algorithms

6 Conclusion

Acquaintance Network

- As in the experiment by Milgram
- . . . or as given by social networking sites such as Facebook, LinkedIn. . .

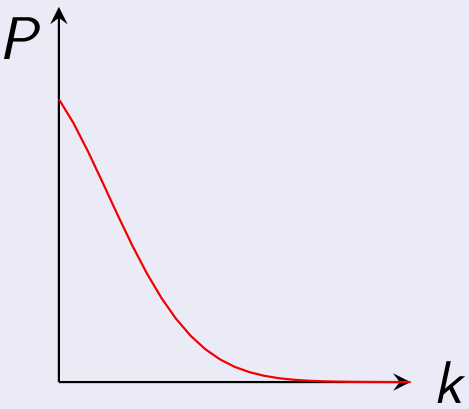


Acquaintance Network

- As in the experiment by Milgram
- . . . or as given by social networking sites such as Facebook, Li

Network characteristics

- Logarithmic typical distance
- Strong local clustering
- Gaussian degree distribution [Amaral et al., 2000]



Actors and Scientists Networks

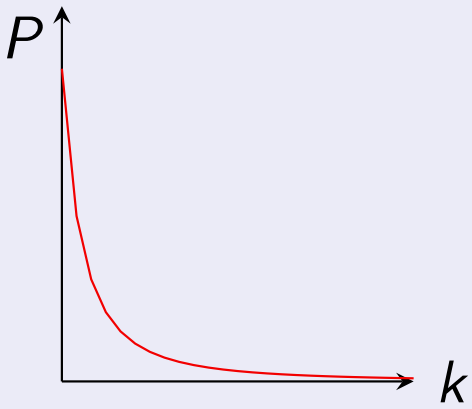
- Bipartite graphs Actor-Movie and Scientist-Publication
- Corresponding undirected graphs:
  - actors appearing in the same movie
  - scientists who coauthored a paper
- Bacon's/Erdős number: distance in the graph to the corresponding vertex

Actors and Scientists Networks

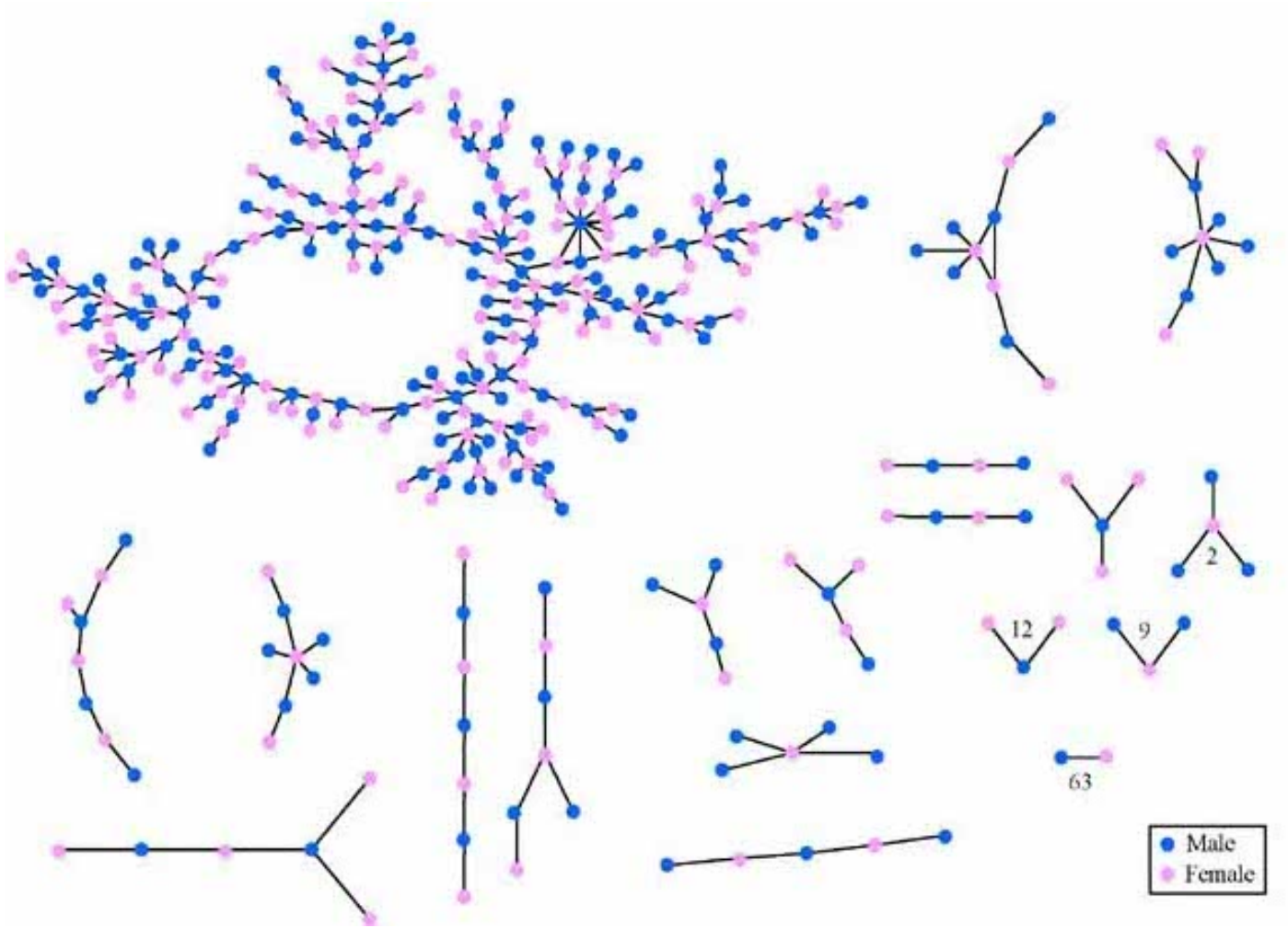
- Bi
- Co
- Ba ing ve

Network characteristics

- Logarithmic typical distance
- Strong local clustering
- Power-law degree distribution ($2 \leq \gamma \leq 3$), with a possible tail cutoff [Amaral et al., 2000]



Sex Networks

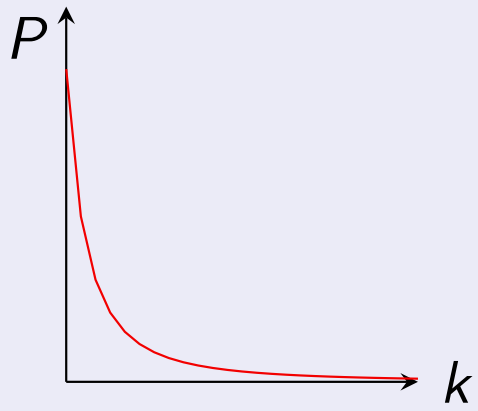


[Amaral et al., 2000]

Sex N

Network characteristics

- In this particular case (small and incomplete community): [Amaral et al., 2000]
  - Unconnected network, long typical distance
  - No local clustering (the graph is almost bipartite!)
- But for larger studies [Liljeros et al., 2001]:
  - Logarithmic typical distance
  - No strict local clustering because of predominance of heterosexuality, but some kind of locality
  - Power-law degree distribution ($\gamma \approx 2.5$ for females, $\gamma \approx 2.3$ for males)



al., 2000]

Sociological aspects

- An individual has two kinds of social network:
  - his actual connections with individuals (acquaintances, relationships, etc.) (explicit network)
  - his virtual connections with similar individuals (implicit network)
- Sociologically speaking, four kinds of connections between individuals [Smith et al., 2007], depending on the kind of social capital [Lin, 2001] considered:

		Implicit link: Yes	Implicit link: No
Explicit link	Yes	Actual bonding	Actual bridging
	No	Potential bonding	Potential bridging

- Both bonding and bridging necessary to make one's life successful

Outline

1 Introduction

2 Basics of Graph Theory

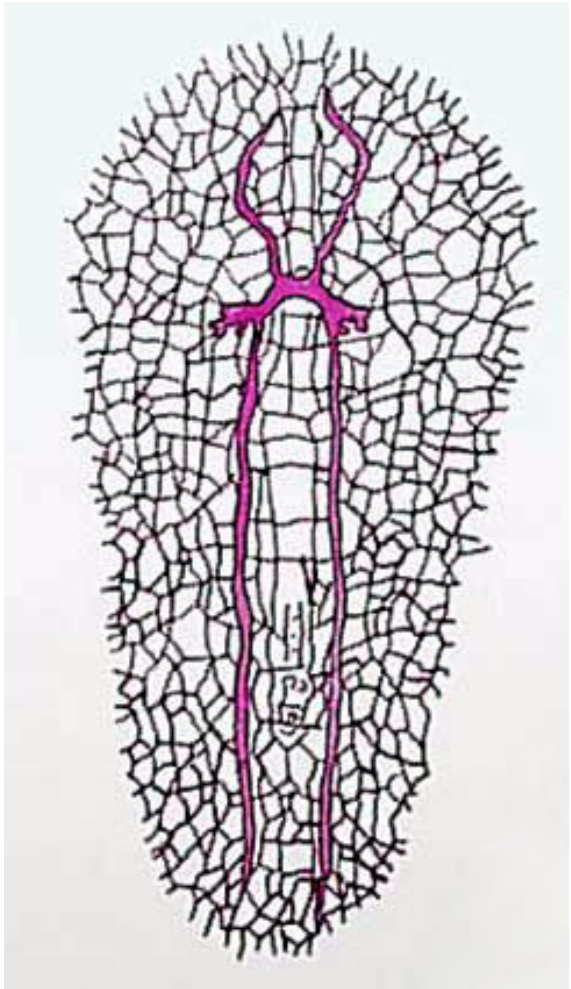
3 Characteristics of Real-World Networks
- Social Networks
- Natural Networks
- Artificial Networks

4 Models of Networks

5 Graph Mining Algorithms

6 Conclusion

Neural Networks

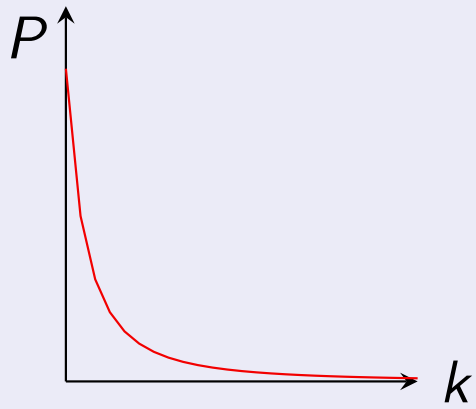


(Dorling Kindersley, dkimages)

Neural Networks

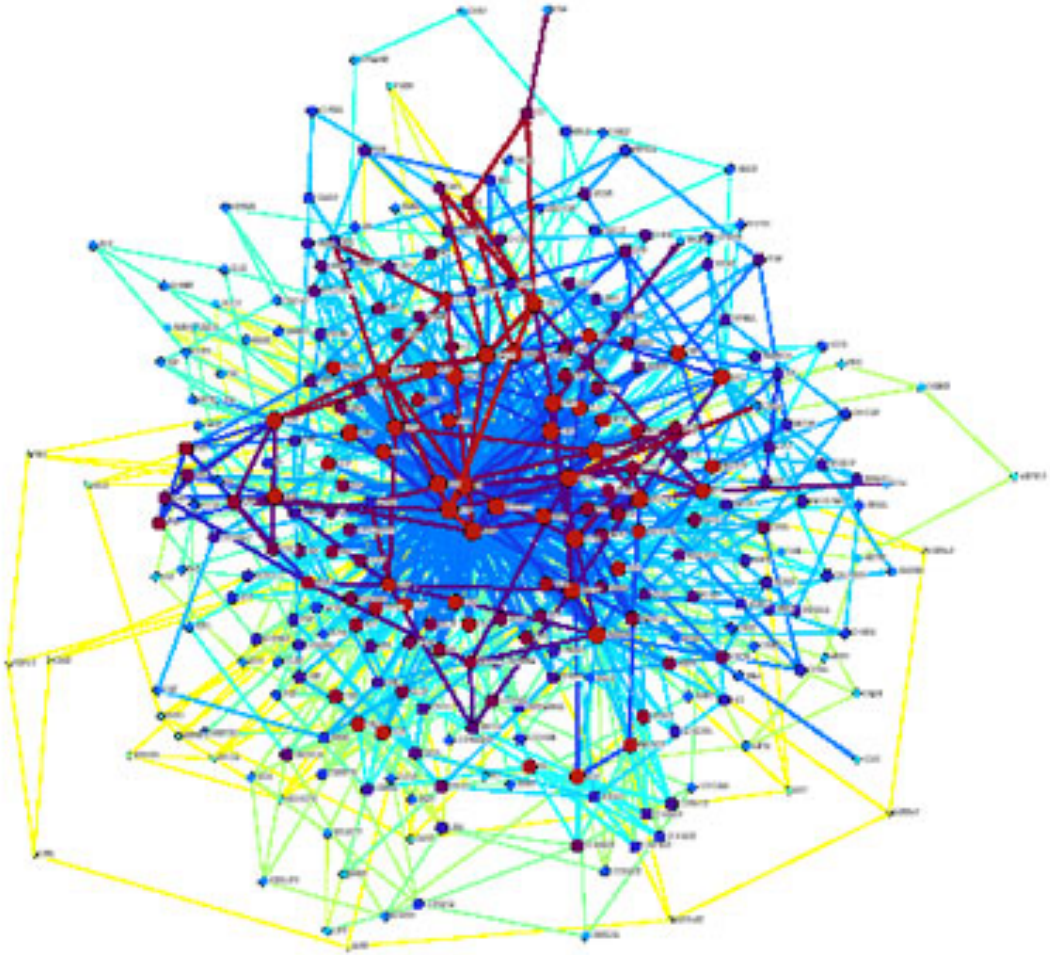
Network characteristics

- Logarithmic typical distance [Watts and Strogatz, 1998]
- Strong local clustering
- Power-law degree distribution



(Dorling Kindersley, dkimages)

Metabolic Networks

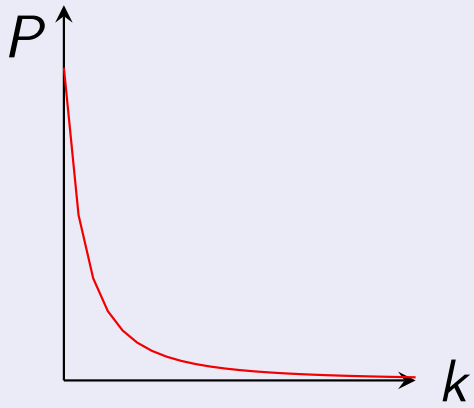


(Laboratory of Computer Engineering, Technical University of Helsinki)

Metabolic Networks

Network characteristics

- Logarithmic typical distance
- Strong local clustering
- Power-law degree distribution ($2 \leq \gamma \leq 2.4$) [Jeong et al., 2000]



(Laboratory of Computer Engineering, Technical University of Helsinki)

Outline

1 Introduction

2 Basics of Graph Theory

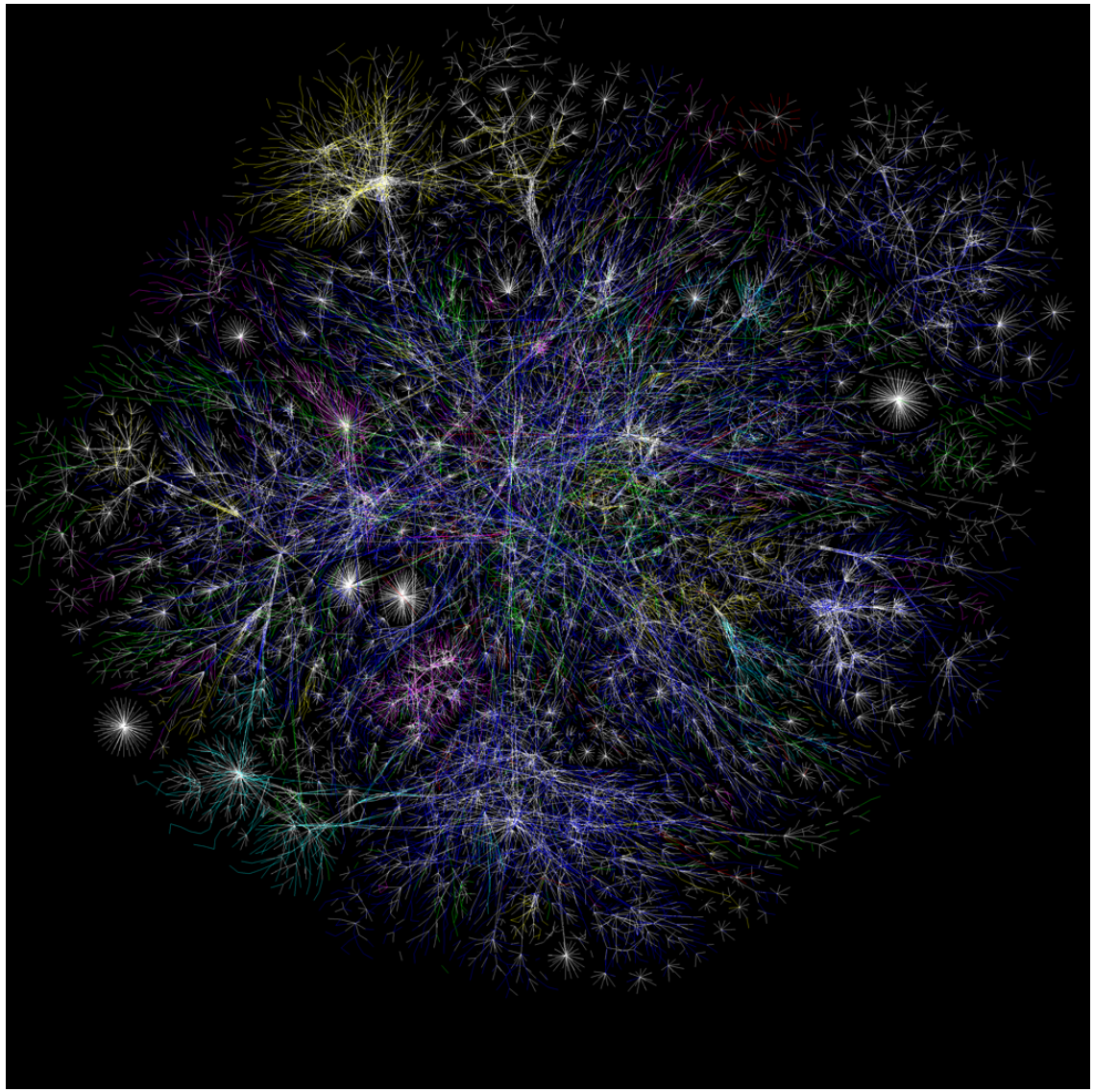
3 Characteristics of Real-World Networks
- Social Networks
- Natural Networks
- Artificial Networks

4 Models of Networks

5 Graph Mining Algorithms

6 Conclusion

The Internet: physical connections between LANs

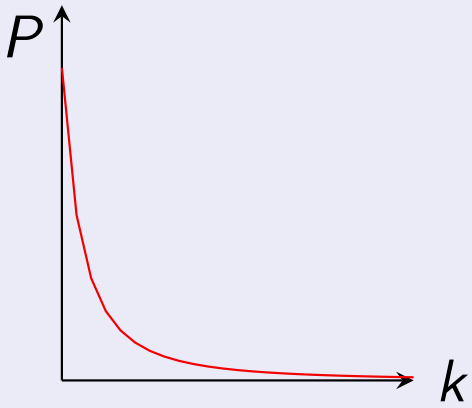


http://www.opte.org/

The Internet: physical connections between LANs

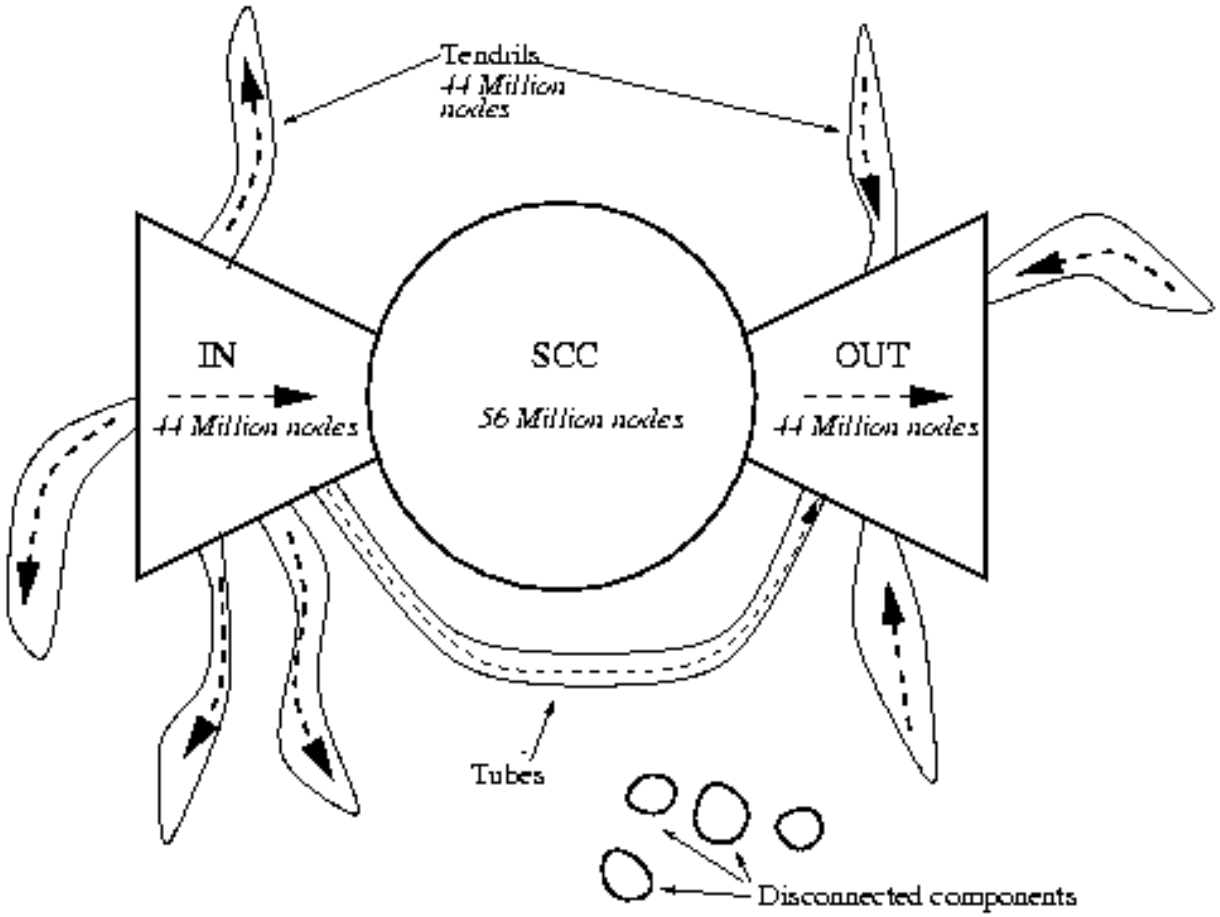
Network characteristics

- Logarithmic typical distance
- Strong local clustering
- Power-law degree distribution ($\gamma \approx 2.2$) [Faloutsos et al., 1999]



http://www.opte.org/

The Web: logical hyperlinks between Web pages

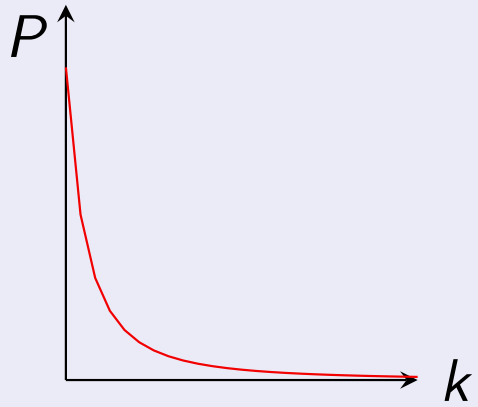


[Broder et al., 2000]

The Web: logical hyperlinks between Web pages

Network characteristics

- Directed graph
- Logarithmic typical distance
- Strong local clustering
- Power-law indegree and outdegree distribution ($2 \leq \gamma \leq 3$) [Broder et al., 2000]



[Broder et al., 2000]

Scientific Citations Network

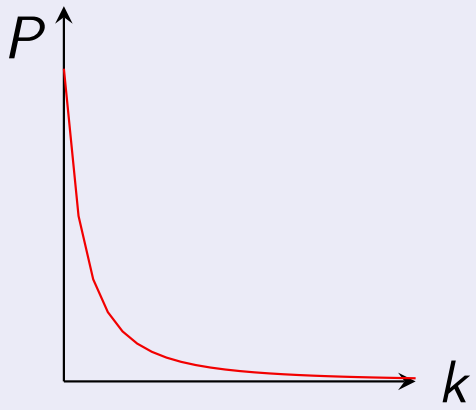
Vertices: Scientific publications

Edges: Citation links

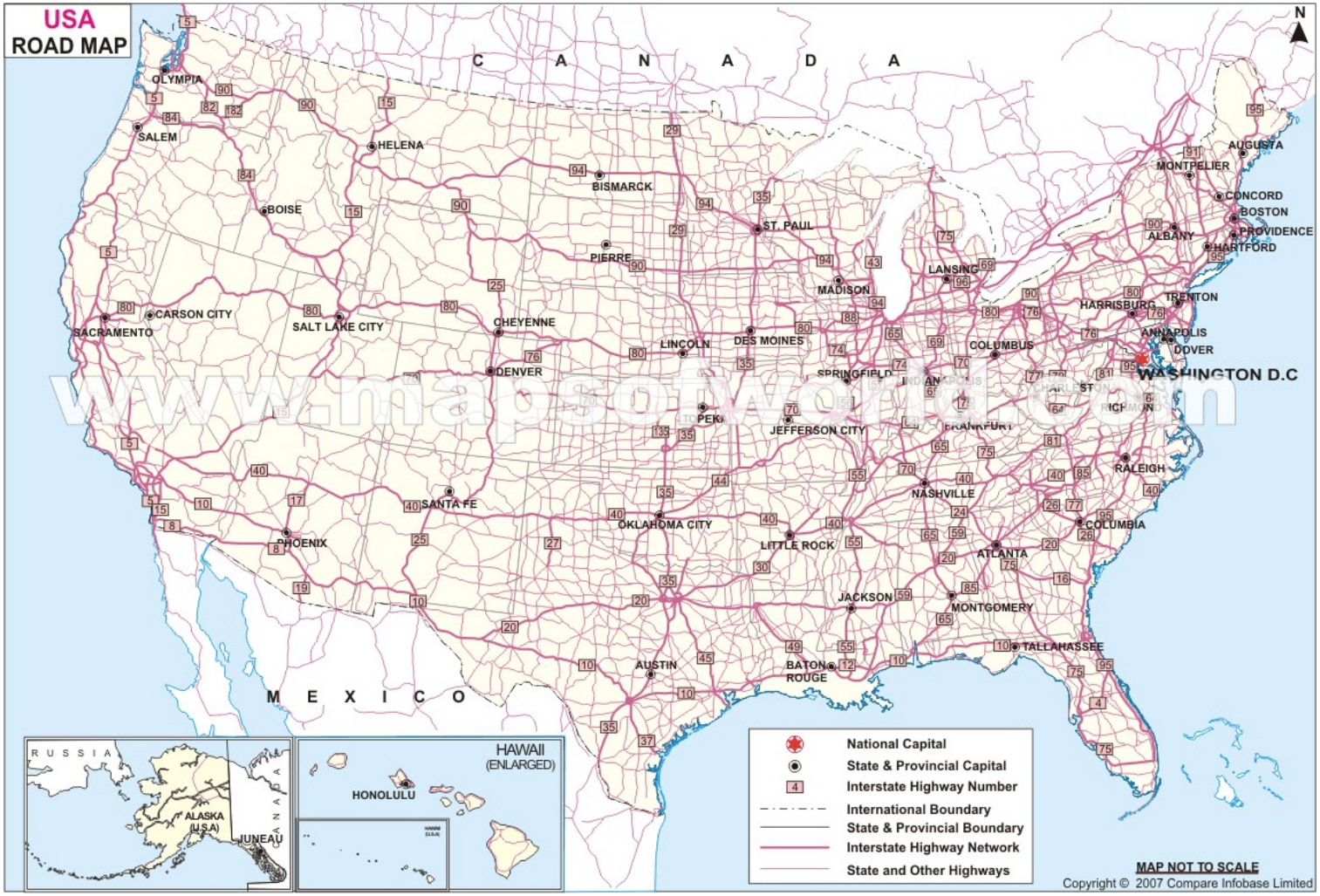
Scientific Citations Network

Network characteristics

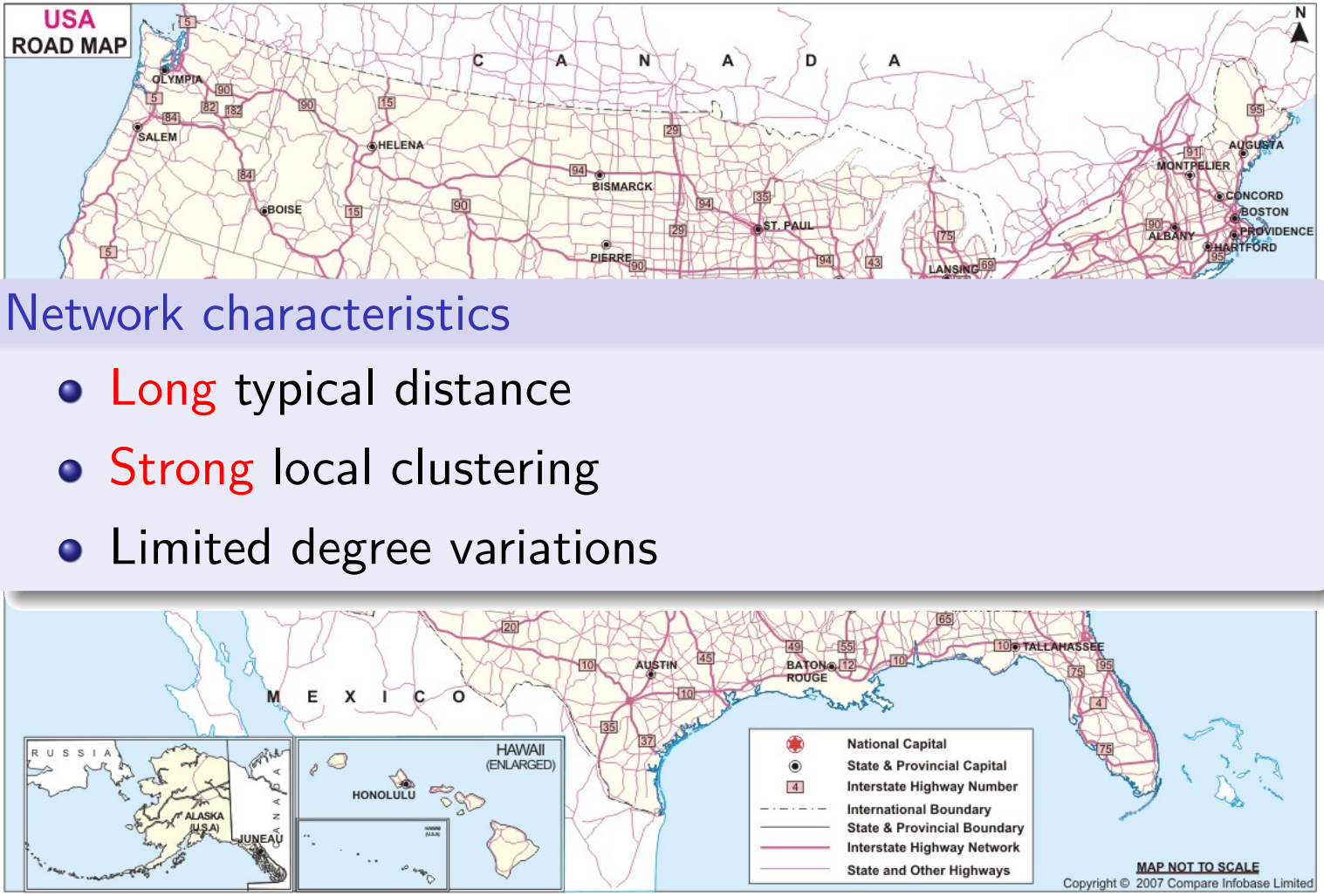
- Directed graph
- No cycles! No strong connectivity.
- Strong local clustering (on the underlying undirected graph)
- Power-law indegree and outdegree distribution ($2 \leq \gamma \leq 3$)



Transportation Networks



Transportation Networks



Network characteristics

- Long typical distance
- Strong local clustering
- Limited degree variations

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks

4 Models of Networks
- Random Networks
- Small Worlds
- Scale-Free Networks

5 Graph Mining Algorithms

6 Conclusion

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks

4 Models of Networks
- Random Networks
- Small Worlds
- Scale-Free Networks

5 Graph Mining Algorithms

6 Conclusion

Random Networks [Solomonoff and Rapoport, 1951, Erdős and Rényi, 1960]

Construction

1. Start with n vertices and a probability p.
2. For each pair of vertices (u, v), insert an edge between u and v with probability p.

Random Networks [Solomonoff and Rapoport, 1951, Erdős and Rényi, 1960]

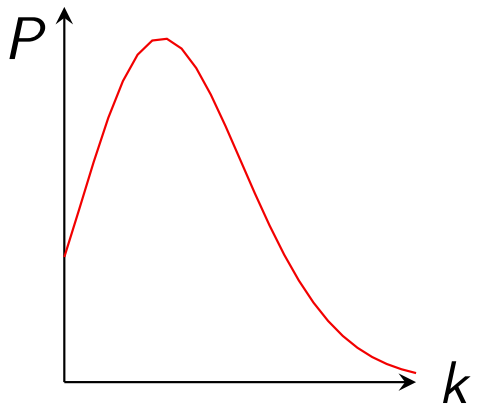
Construction

1. Start with n vertices and a probability p.
2. For each pair of vertices (u, v), insert an edge between u and v with probability p.

- Sparse if $p \ll 1$
- Logarithmic typical distance (inside the giant connected component)!
- No local clustering.

Degree distribution in random networks

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k} \sim \frac{(pn)^k e^{-pn}}{k!}$$



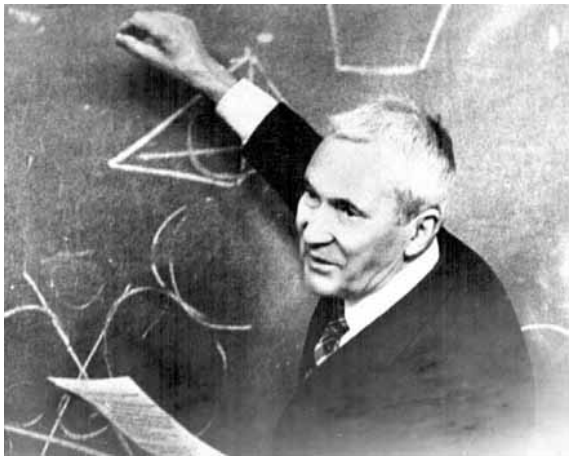
Exercise

Prove this.

Remark

One can construct random graphs with an arbitrary degree distribution (more complicated); still no local clustering, obviously.

Kolmogorov's 0-1 Law



- Andrey Kolmogorov (1903-1987): Brilliant mathematician, laid the foundations of modern probability theory, between other very varied and important works, with impact in mathematics, physics, and computer science.

- Kolmogorov's 0-1 Law: Given an infinite sequence of independent random variables, events that are probabilistically independent of any finite subset of the random variables have probability either 0 or 1.
- Application: The probability that there is a (single) giant connected component goes to 1 when the size of the graph goes either to 0 or to $+\infty$, depending on p (actually, the transition is $p \geq 1/n$).

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks

4 Models of Networks
- Random Networks
- Small Worlds
- Scale-Free Networks

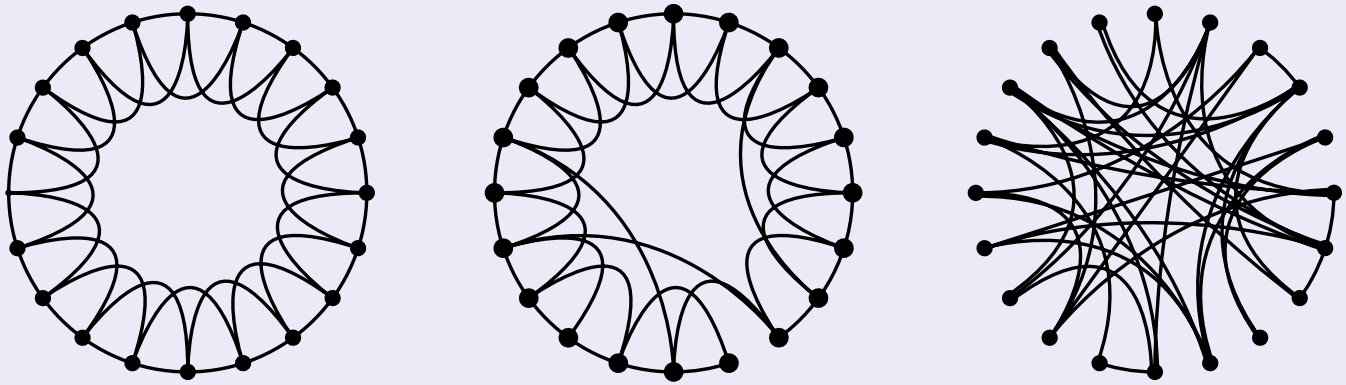
5 Graph Mining Algorithms

6 Conclusion

Small Worlds [Watts and Strogatz, 1998, Watts, 1999]

Construction

1. Start with a regular lattice (a grid).
2. With probability p, reroute each edge randomly.

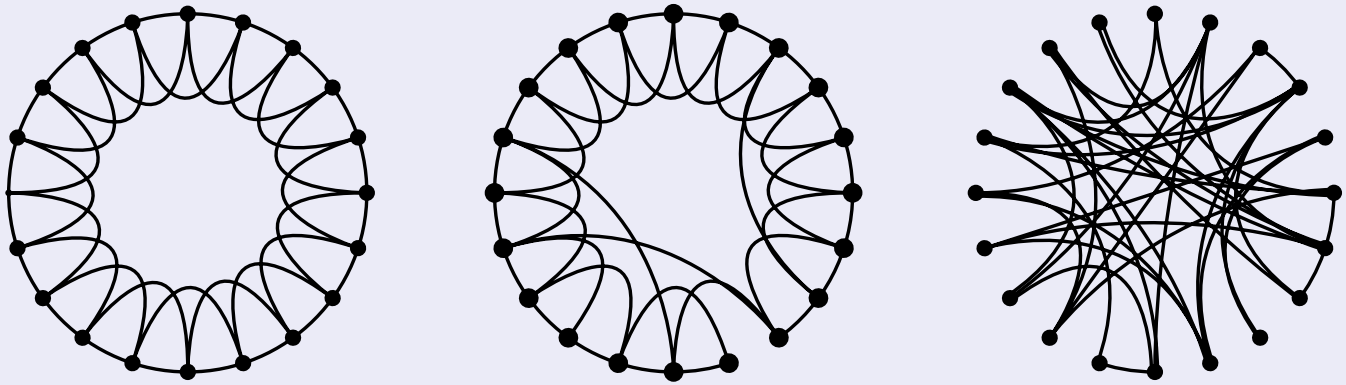


[Watts and Strogatz, 1998]

Small Worlds [Watts and Strogatz, 1998, Watts, 1999]

Construction

1. Start with a regular lattice (a grid).
2. With probability p, reroute each edge randomly.

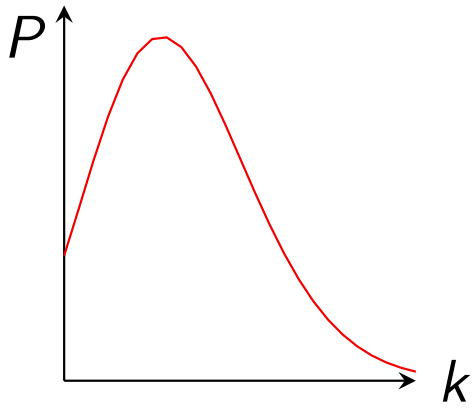


[Watts and Strogatz, 1998]

Sparse.

Characteristics of Small Worlds

- For $p = 0$: lattice (strong local clustering)
- For $p = 1$: random graph (small typical distance)
- Somewhere in between:
  - Small typical distance (thanks to rerouting)
  - Strong local clustering (thanks to the initial lattice)
  - Degree distribution resembling a Poisson.



Measuring the local clustering

$$C_G = \frac{3 \times (\text{number of triangles in } G)}{\text{number of connected triples in } G}$$

- $C_{\text{fg}} = 1$ for a fully connected graph
- $C_{\text{rg}} = p$ for a random graph
- A graph G has strong local clustering if $C_G \gg C_{\text{rg}}$ (for the random graph with the same number of edges)

Exercise

Prove $C_{\text{rg}} = p$.

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks

4 Models of Networks
- Random Networks
- Small Worlds
- Scale-Free Networks

5 Graph Mining Algorithms

6 Conclusion

Preferential Attachment [Barabási and Albert, 1999]

Construction

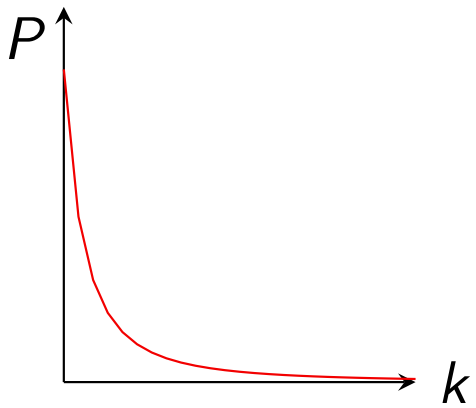
1. Start with a small graph of size m_0, let m be a constant with $m < m_0$.
2. One after the other, $n - m_0$ vertices are added to the graph, connecting them to m existing vertices; the probability of connecting to a vertex is proportional to its degree.

Preferential Attachment [Barabási and Albert, 1999]

Construction

1. Start with a small graph of size m_0, let m be a constant with $m < m_0$.
2. One after the other, $n - m_0$ vertices are added to the graph, connecting them to m existing vertices; the probability of connecting to a vertex is proportional to its degree.

- Sparse if m and n are chosen appropriately.
- Small typical distance.
- Strong local clustering
- Power-law degree distribution (actually, with $\gamma = 3$, but variations allow arbitrary exponents).



Scale-Free Graphs

Graphs with the power-law degree distribution are called scale-free graphs:

- There is no typical scale, or typical order of magnitude for the degree of nodes.
- $$\frac{P(\alpha k)}{P(k)} = \frac{(\alpha k)^{-\gamma}}{k^{-\gamma}} = \alpha^{-\gamma}$$

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks

4 Models of Networks

5 Graph Mining Algorithms
- Search
- Discovery of communities

6 Conclusion

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks

4 Models of Networks

5 Graph Mining Algorithms
- Search
- Discovery of communities

6 Conclusion

Google's PageRank [Brin and Page, 1998]

Idea

Important pages are pages pointed to by important pages.

$$\begin{cases} g_{ij} = 0 & \text{if there is no link between page } i \text{ and } j; \\ g_{ij} = \frac{1}{n_i} & \text{otherwise, with } n_i \text{ the number of outgoing links of page } i. \end{cases}$$

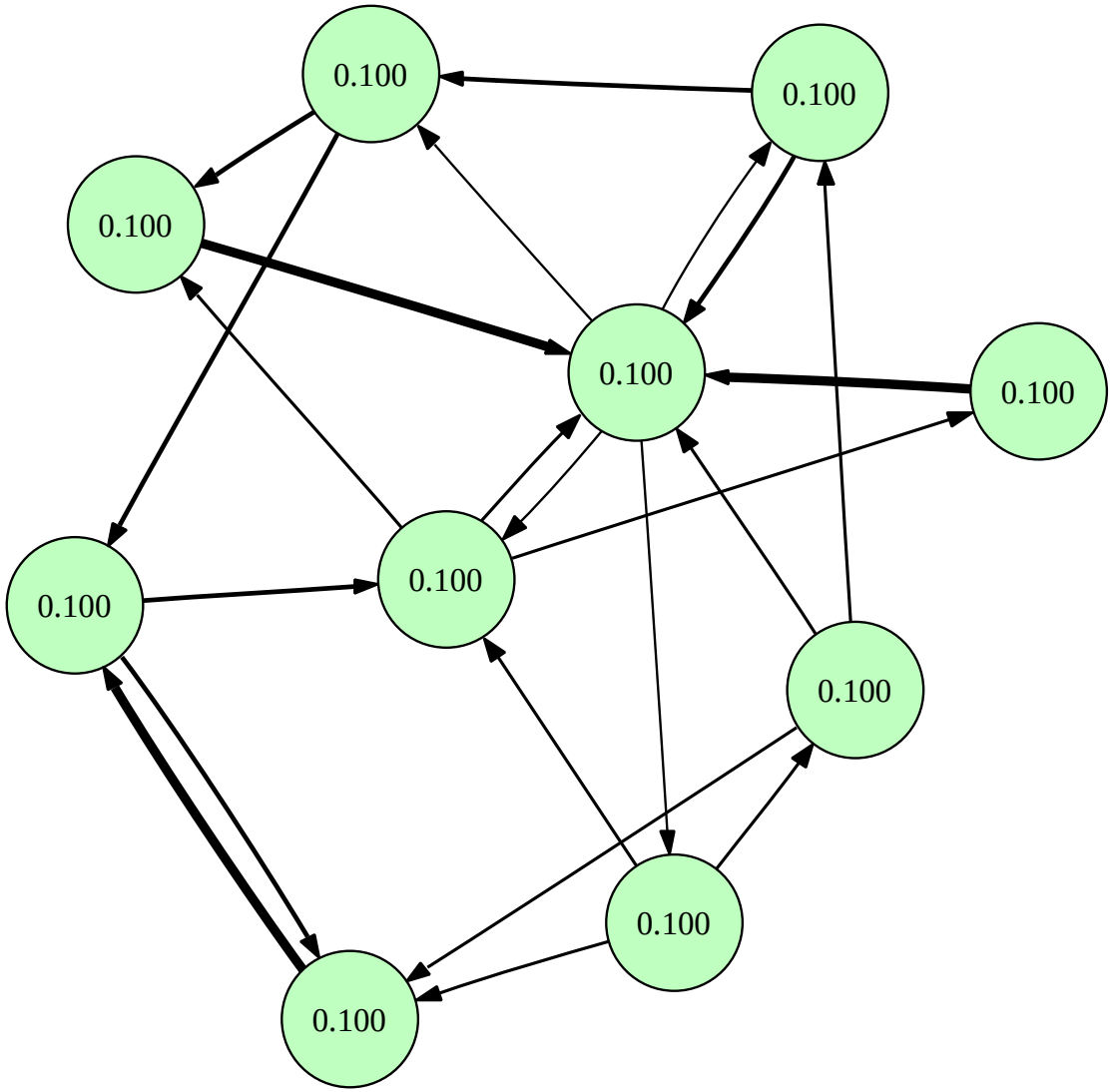
Definition (Tentative)

Probability that the surfer following the random walk in G has arrived on page i at some distant given point in the future.

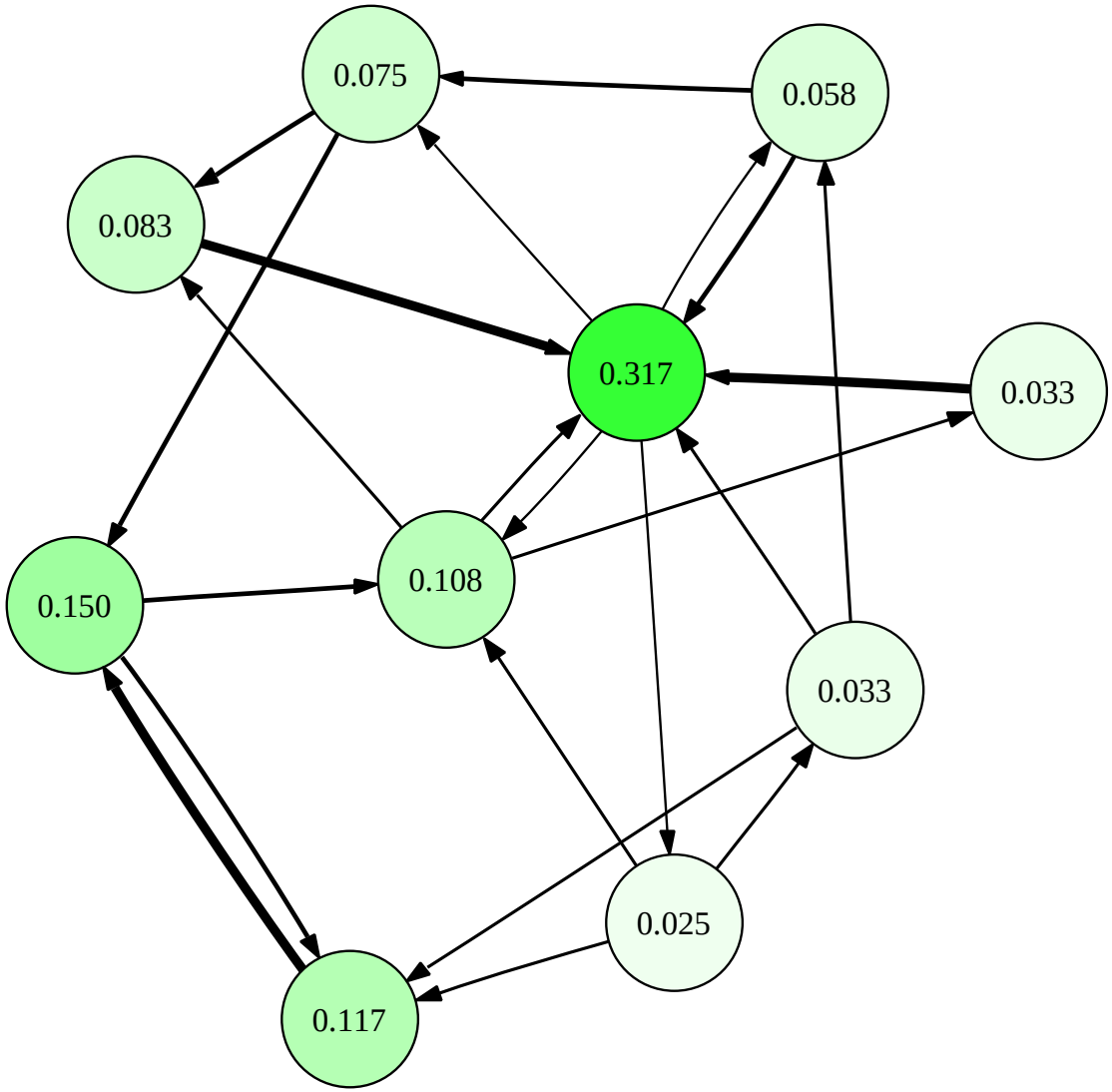
$$\text{pr}(i) = \left(\lim_{k \to +\infty} (G^T)^k v \right)_i$$

where v is some initial column vector.

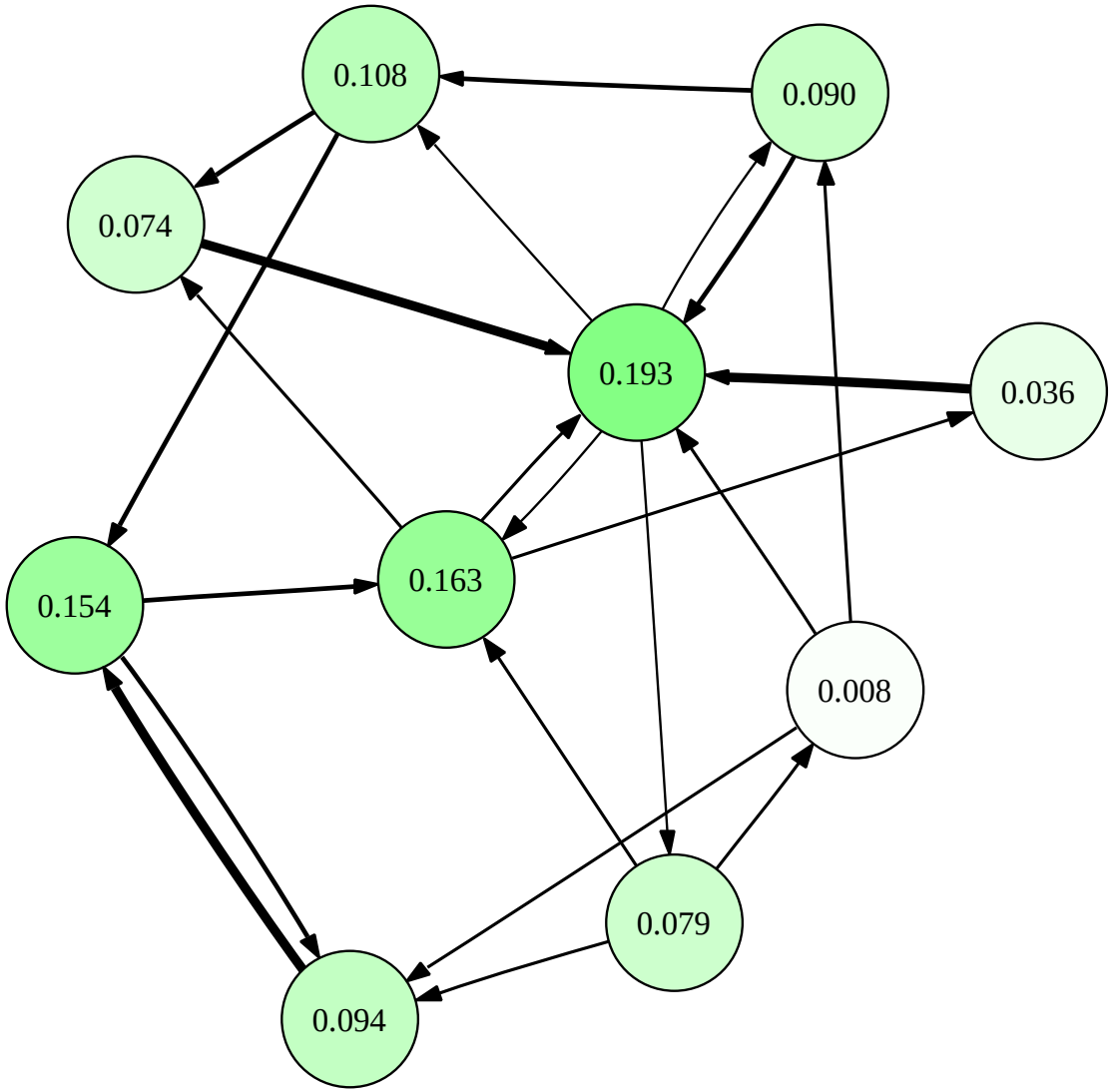
Illustrating PageRank Computation



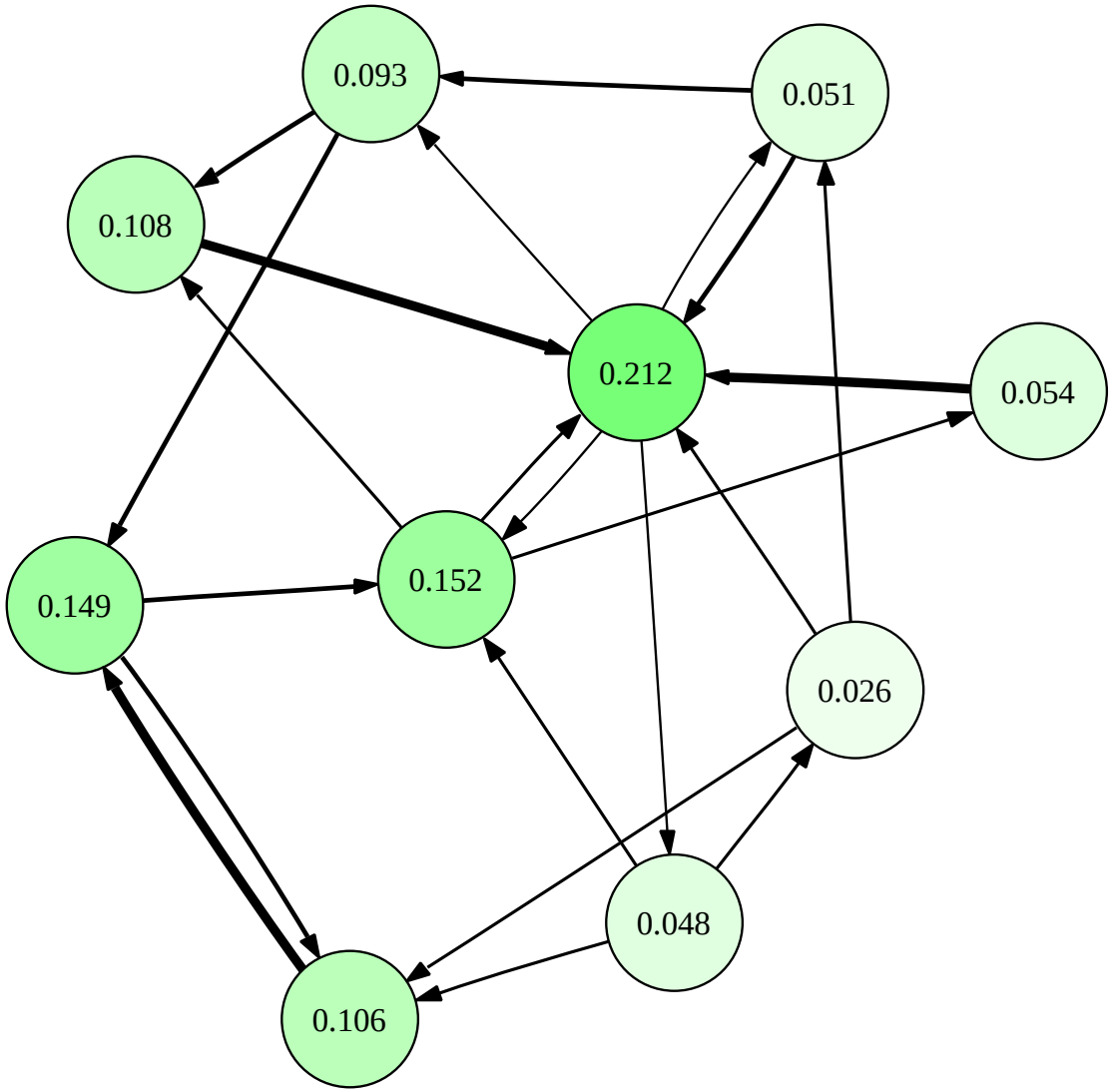
Illustrating PageRank Computation



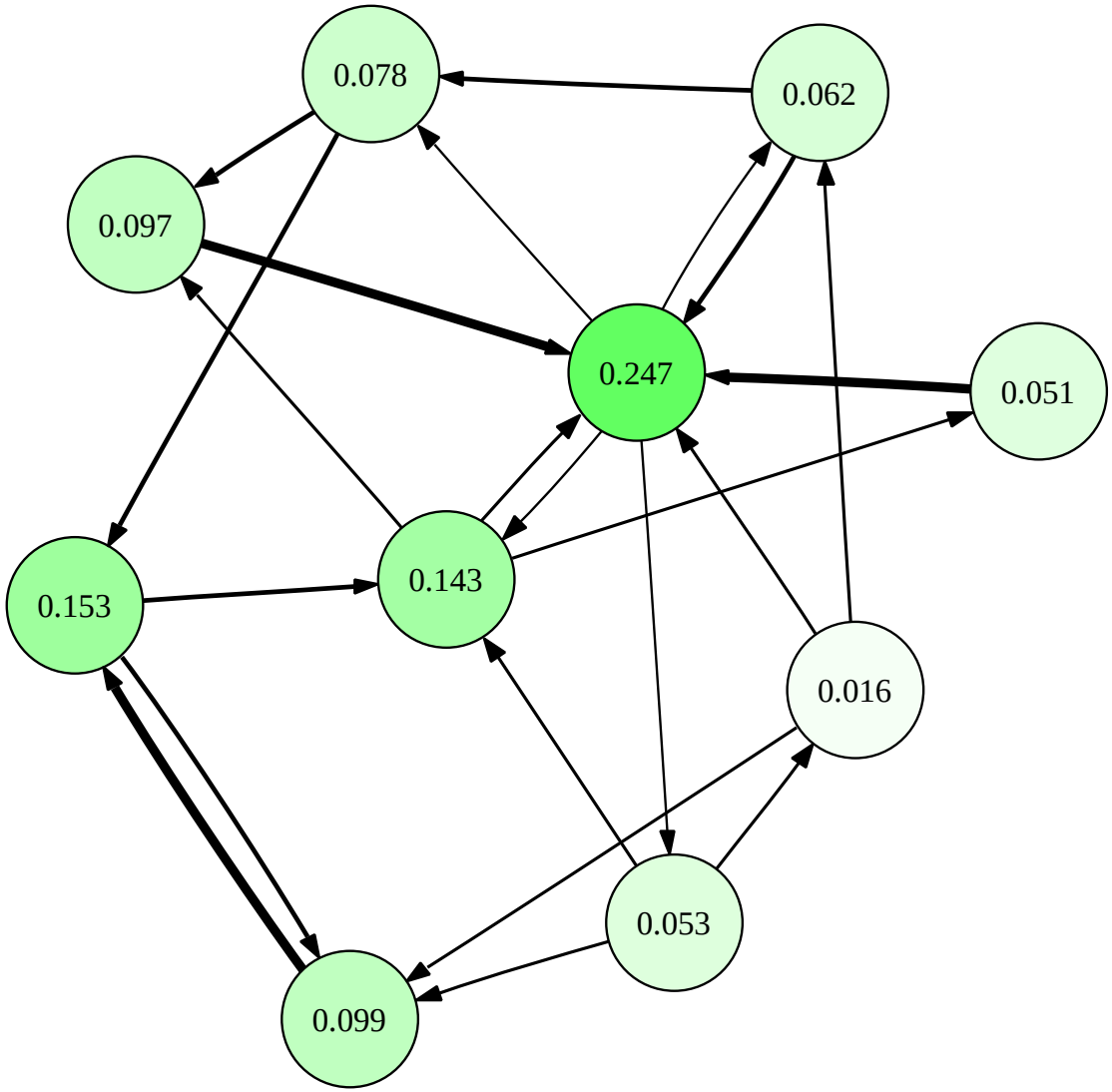
Illustrating PageRank Computation



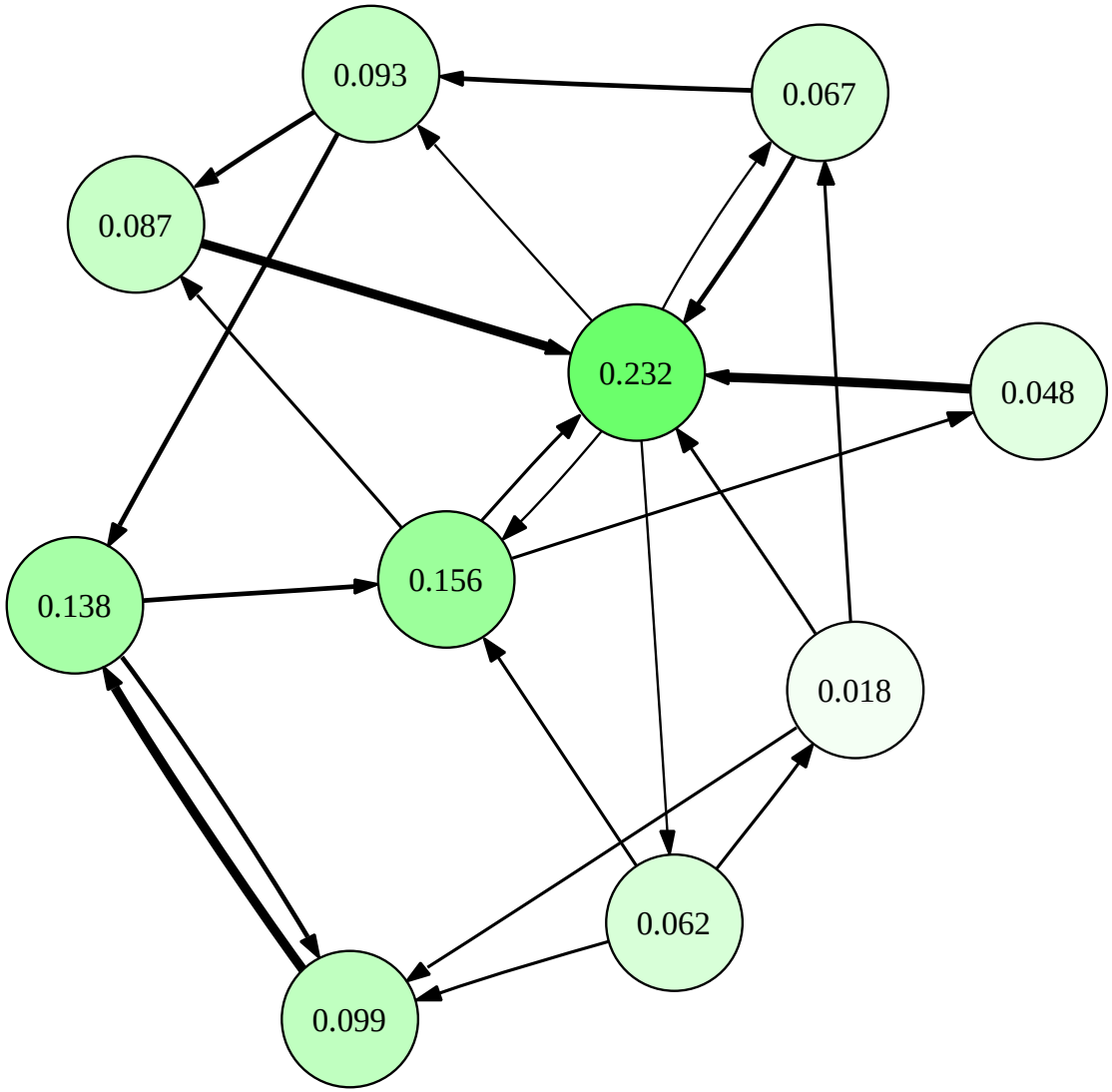
Illustrating PageRank Computation



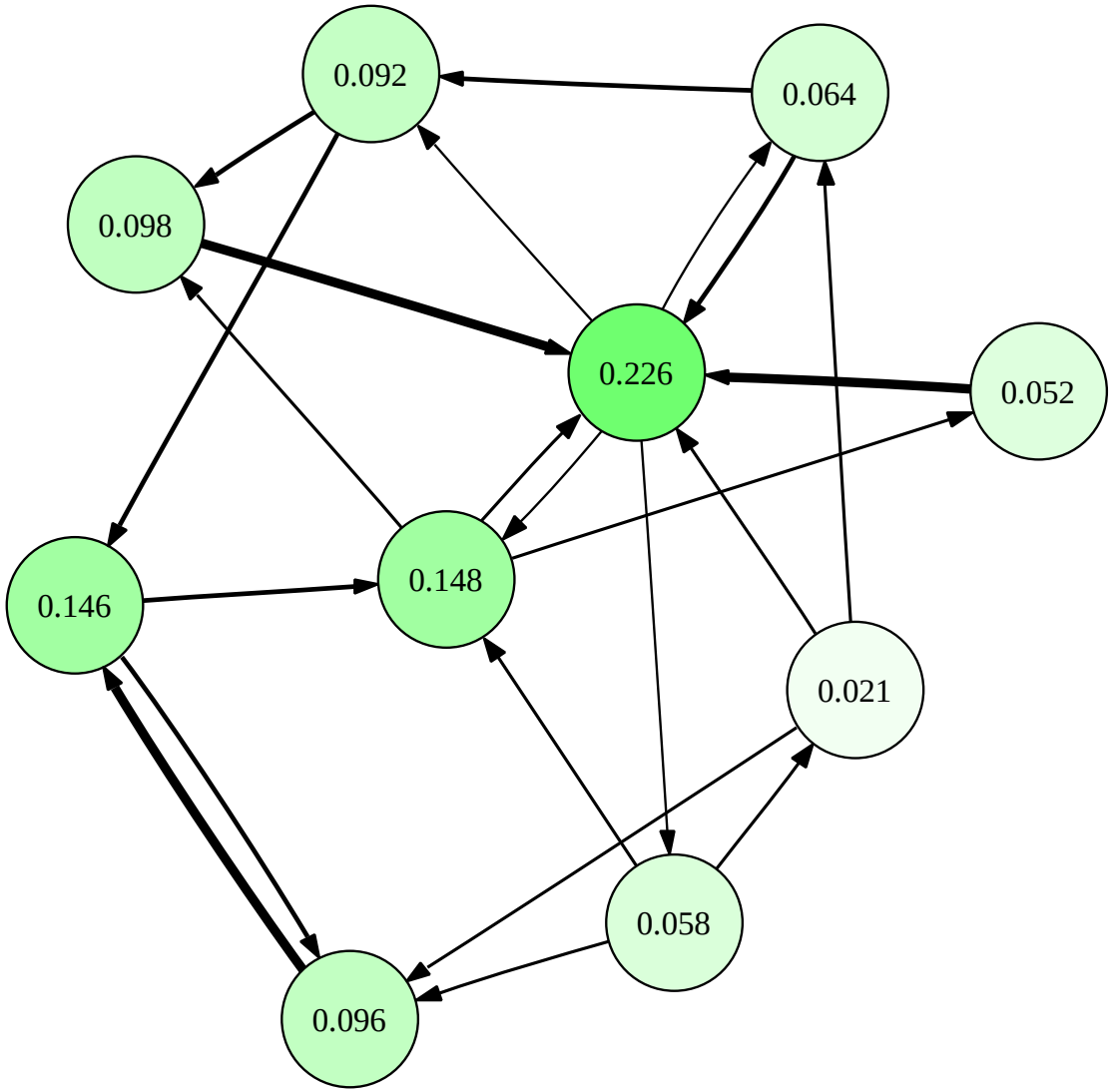
Illustrating PageRank Computation



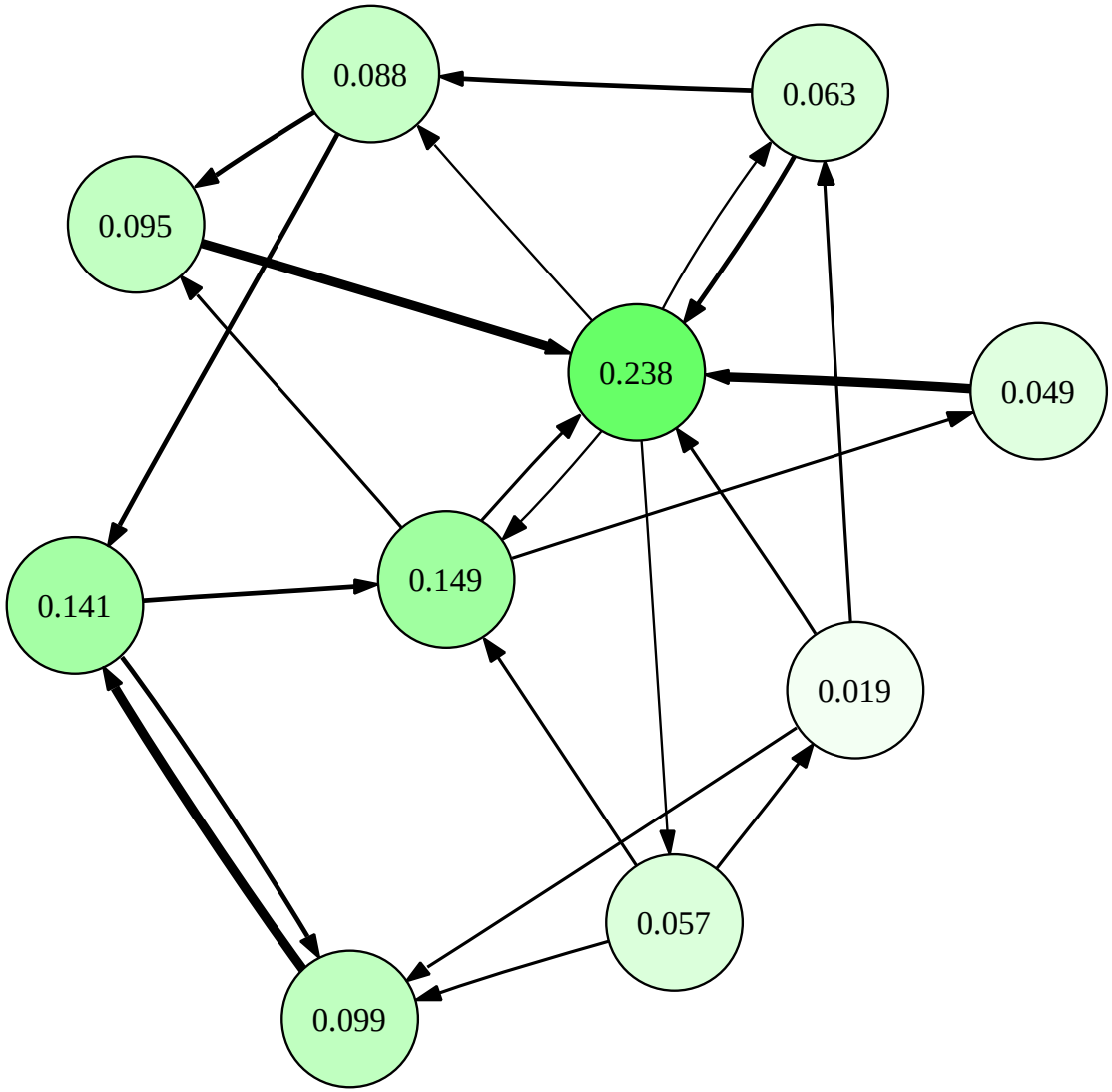
Illustrating PageRank Computation



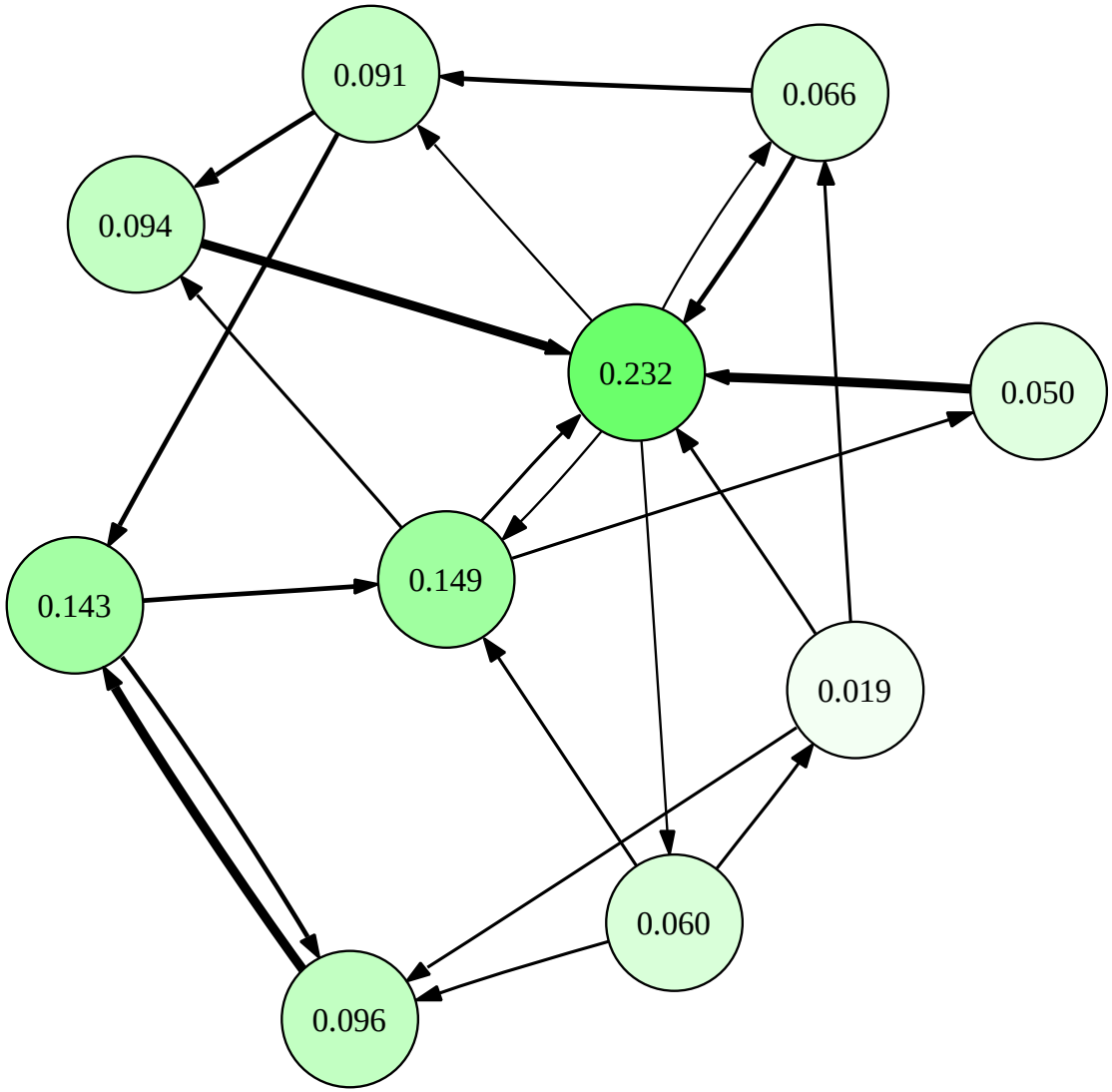
Illustrating PageRank Computation



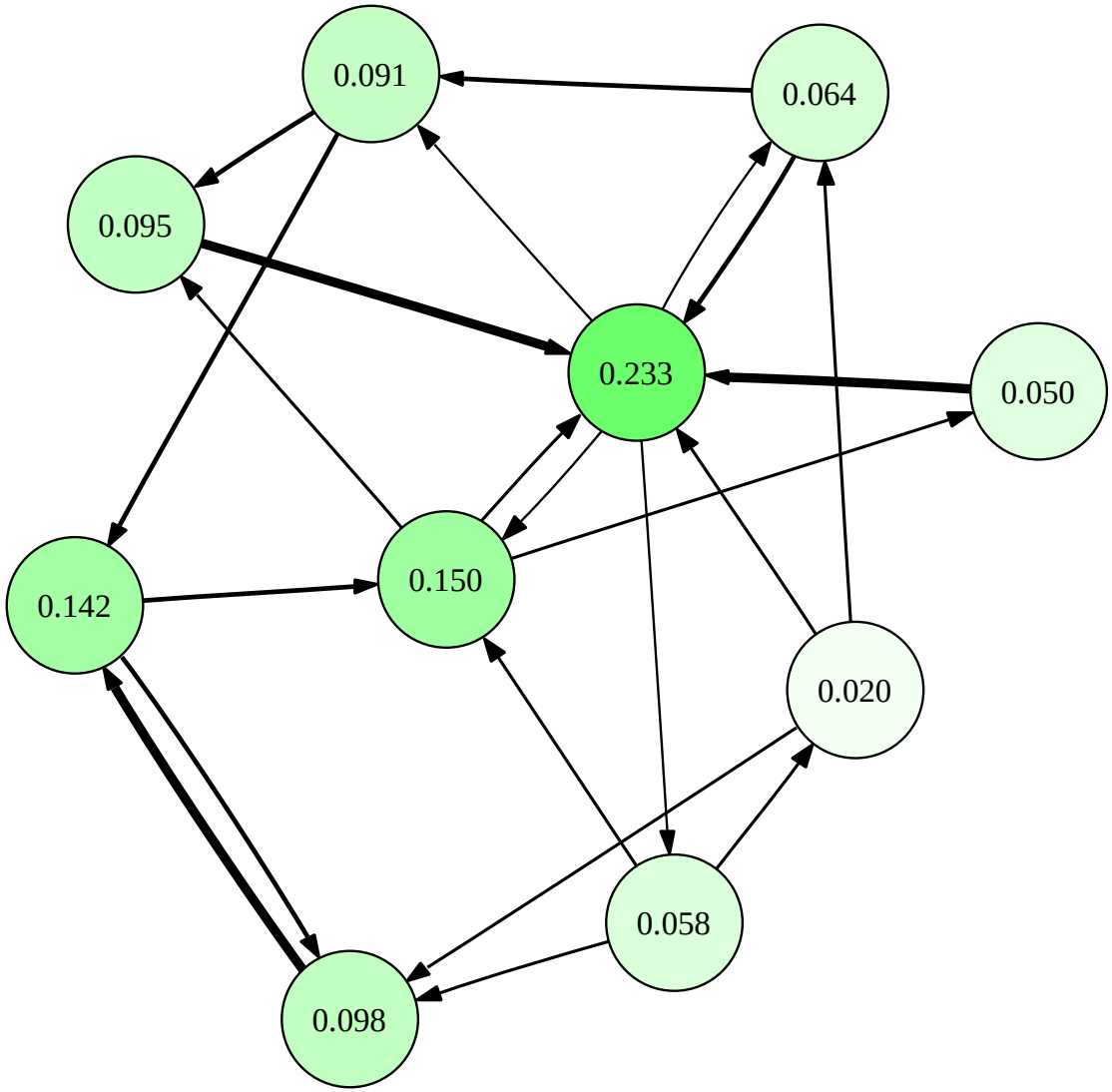
Illustrating PageRank Computation



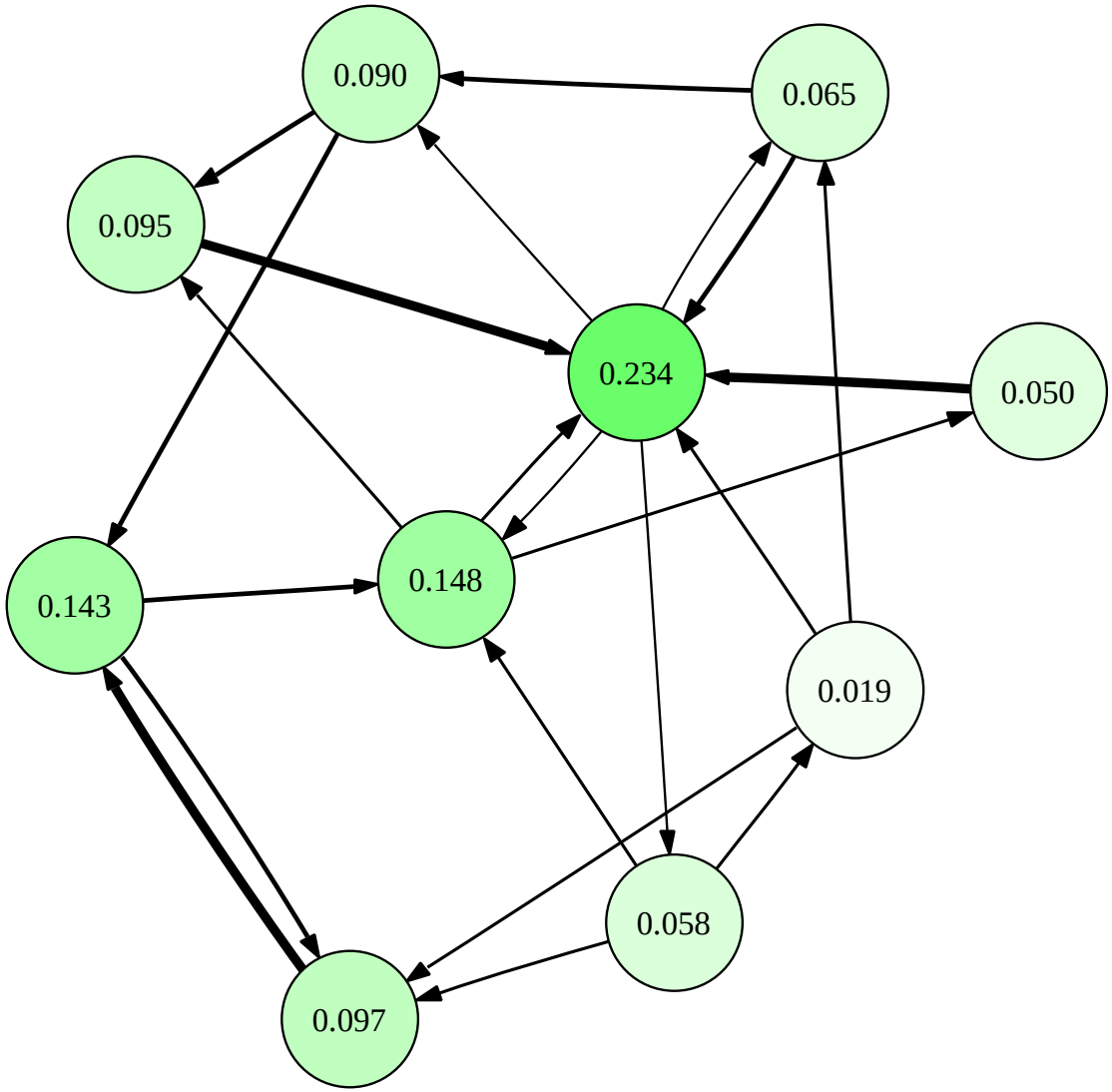
Illustrating PageRank Computation



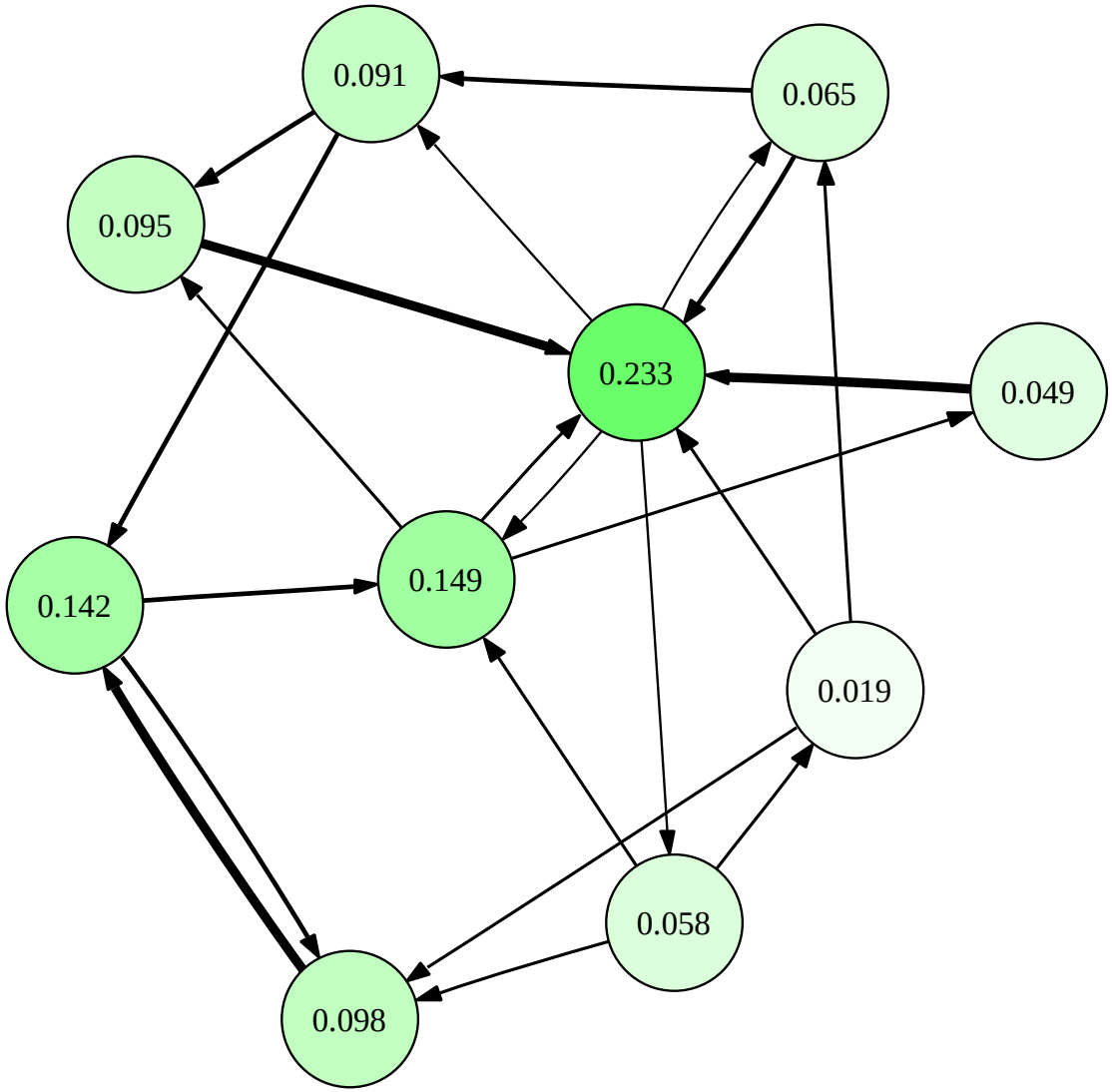
Illustrating PageRank Computation



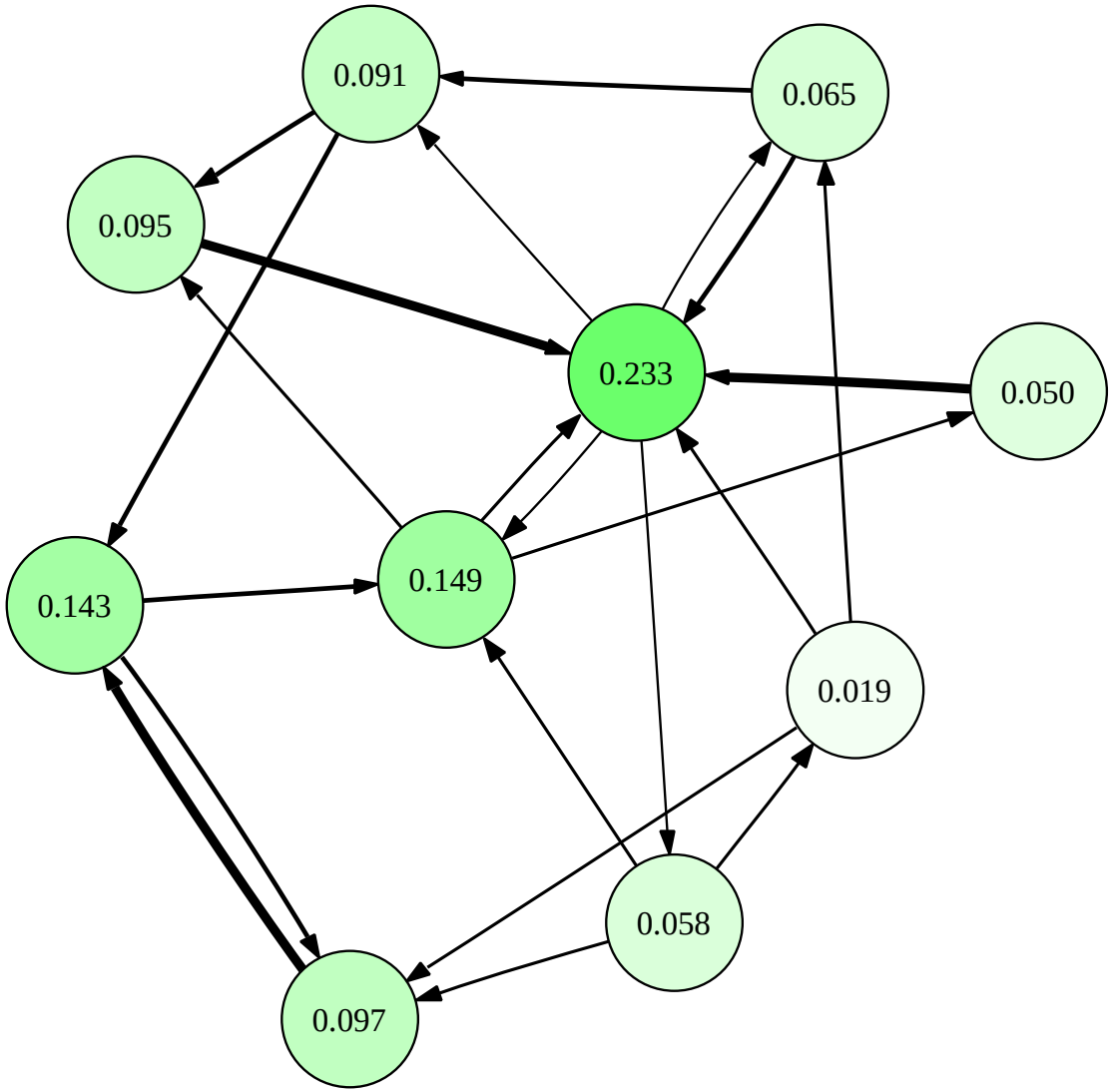
Illustrating PageRank Computation



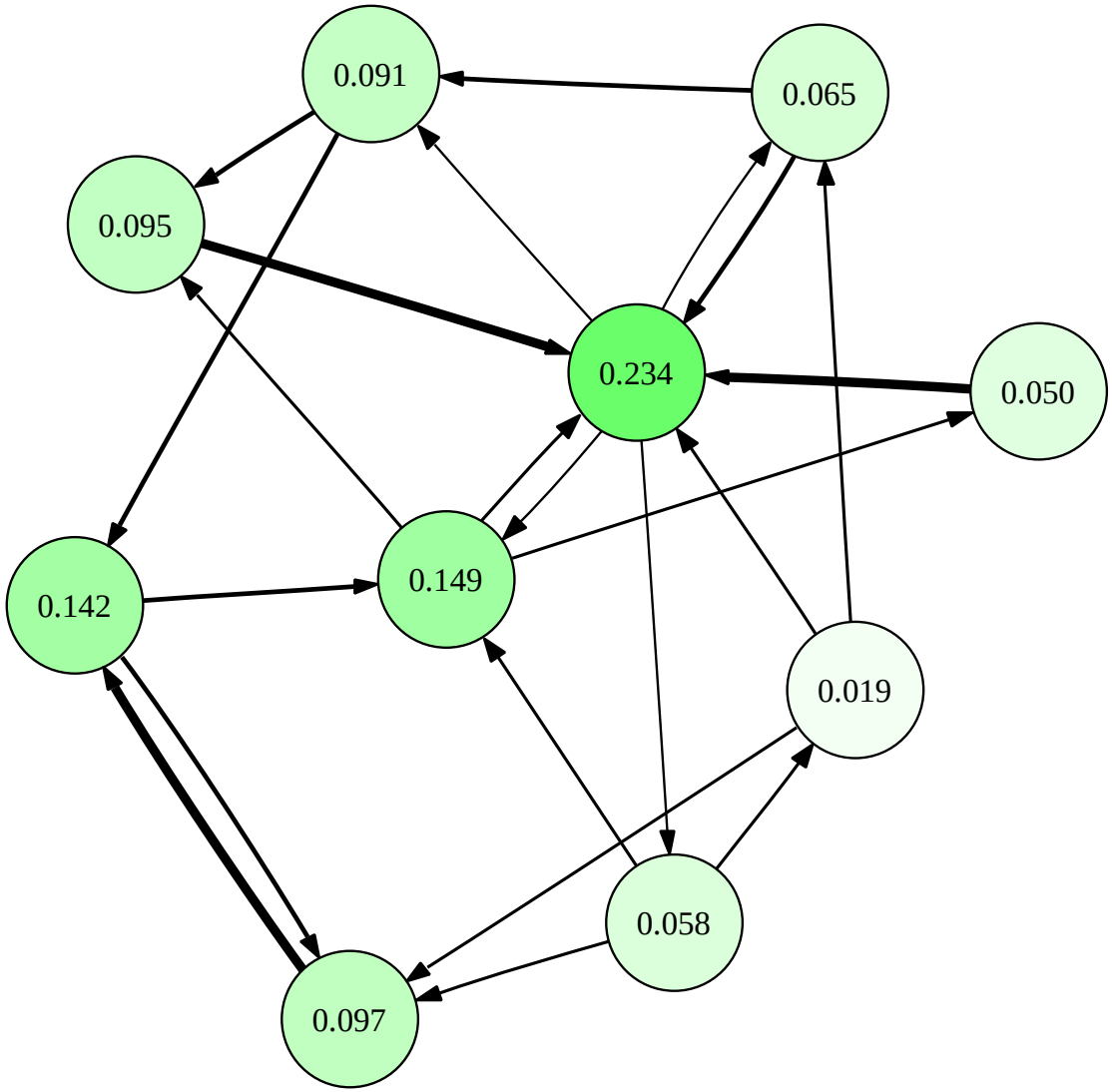
Illustrating PageRank Computation



Illustrating PageRank Computation



Illustrating PageRank Computation



PageRank With Damping

May not always converge, or convergence may not be unique.
To fix this, the random surfer can at each step randomly jump to any page of the Web with some probability d ($1 - d$: damping factor).

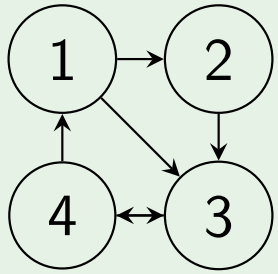
$$\mathrm{pr}(i) = \left(\lim_{k \to +\infty} ((1-d)G^T + dU)^k v \right)_i$$

where U is the matrix with all $\frac{1}{N}$ values with N the number of vertices.

Iterative Computation of PageRank

1. Compute G (often stored as its adjacency list). Make sure lines sum to 1.
2. Let u be the uniform vector of sum 1, $v = u$, w the *zero* vector.
3. While v is different enough from w:
   - Set $w = v$.
   - Set $v = (1 - d)G^T v + du$.

Exercise



Run the first iteration of the PageRank computation.

Using PageRank to Score Search Results

- PageRank: global score, independent of the query
- Can be used to raise the weight of important pages, associated with some scoring function dependent of the query:

$$\text{final}(q, d) = \text{score}(q, d) \times \text{pr}(d),$$

- PageRank only useful in directed graphs! Proportional to degree otherwise

HITS [Kleinberg, 1999]

Idea

Two kinds of important pages: hubs and authorities. Hubs are pages that point to good authorities, whereas authorities are pages that are pointed to by good hubs.

G' adjacency matrix (with 0 and 1 values) of a subgraph of the Web. We use the following iterative process (starting with a and h vectors of norm 1):

$$\begin{cases} a := \frac{1}{\|G'^T h\|} G'^T h \\ h := \frac{1}{\|G' a\|} G' a \end{cases}$$

Converges under some technical assumptions to authority and hub scores.

Using HITS to Order Web Query Results

1. Retrieve the set D of Web pages matching a keyword query.
2. Retrieve the set D^* of Web pages obtained from D by adding all linked pages, as well as all pages linking to pages of D.
3. Build from D^* the corresponding subgraph G' of the Web graph.
4. Compute iteratively hubs and authority scores.
5. Sort documents from D by authority scores.

Less efficient than PageRank, because local scores.

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks

4 Models of Networks

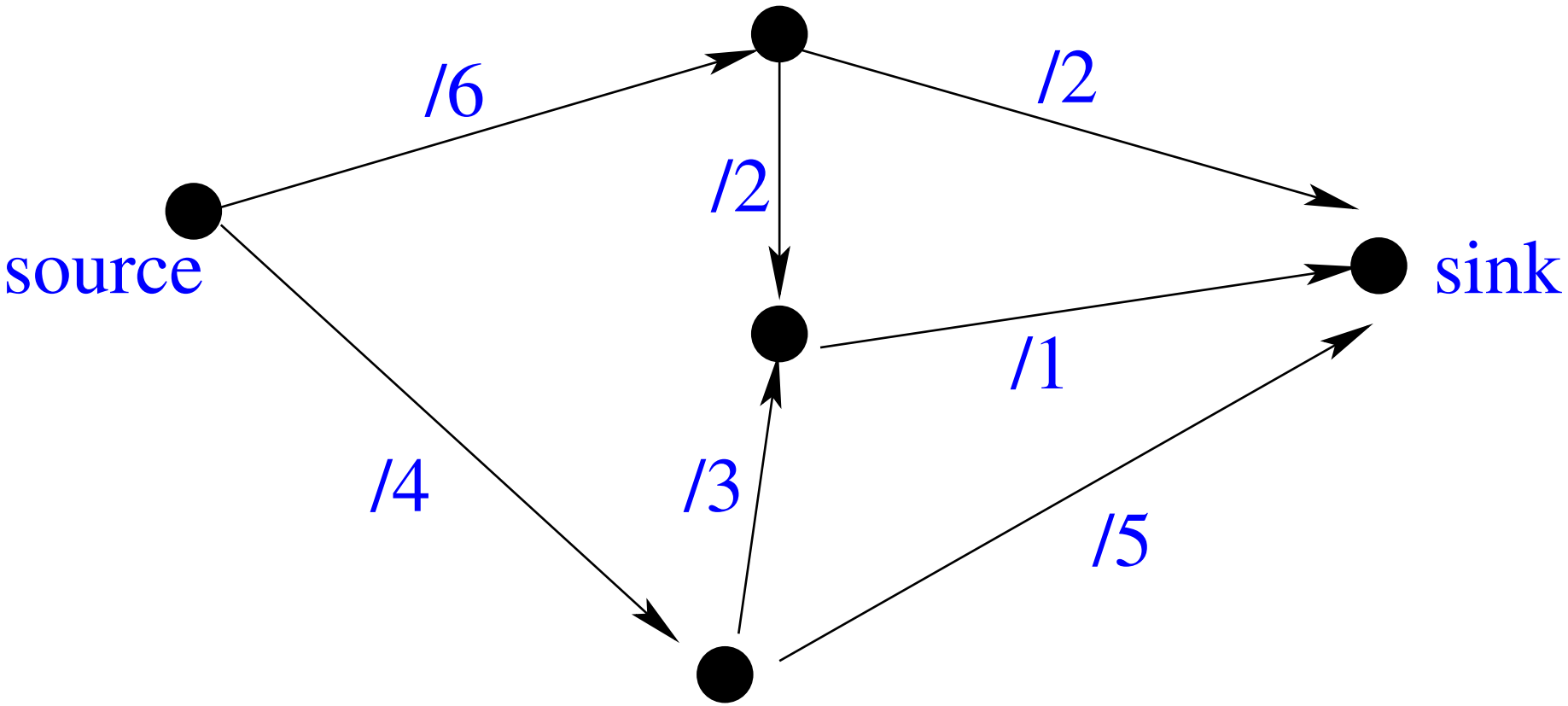
5 Graph Mining Algorithms
- Search
- Discovery of communities

6 Conclusion

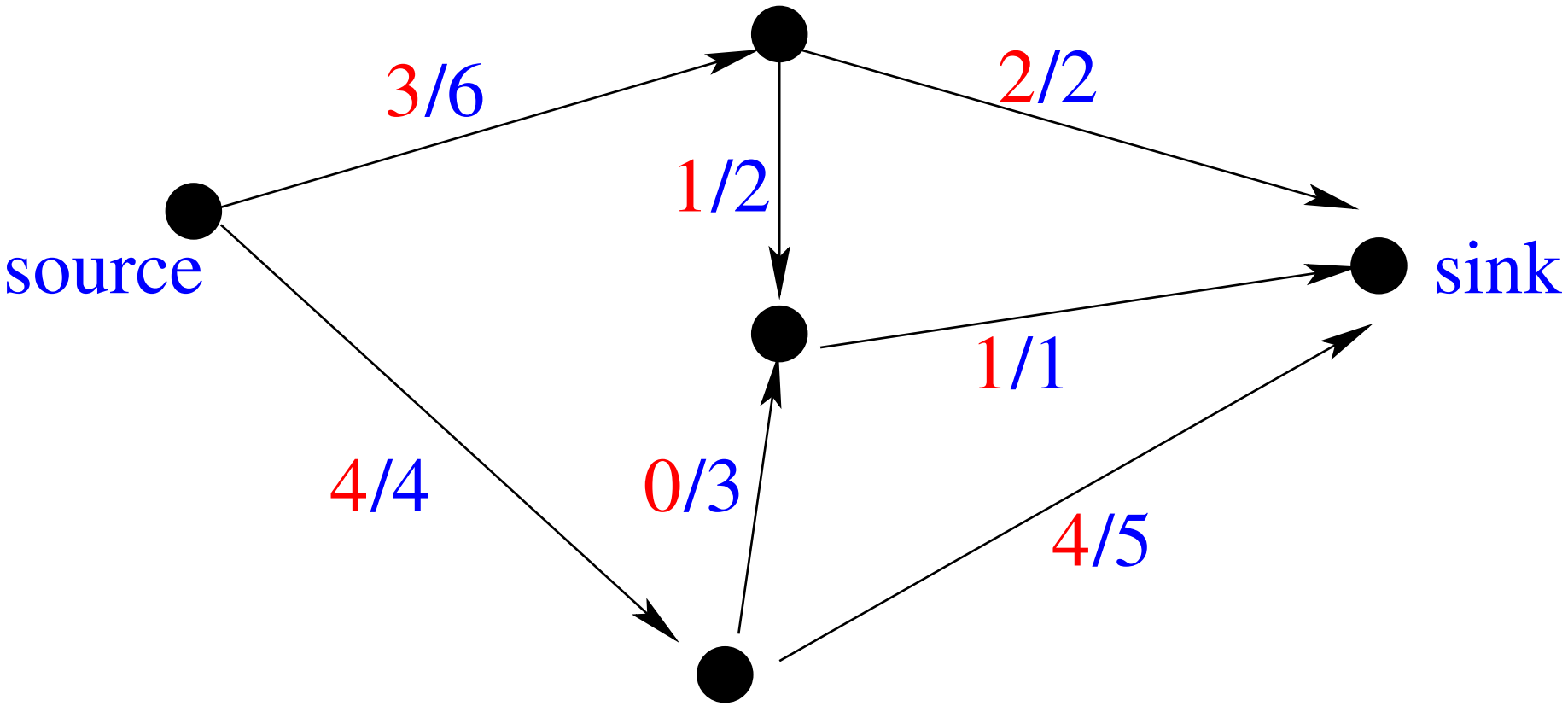
Discovery of communities

- Classical problem in social networks: identifying communities of users (or of content) using the graph structure
- Two subproblems:
  1. Given some initial vertex or vertex set, finding the corresponding community
  2. Given the graph as a whole, finding a partition in communities

Maximum Flow / Minimum Cut

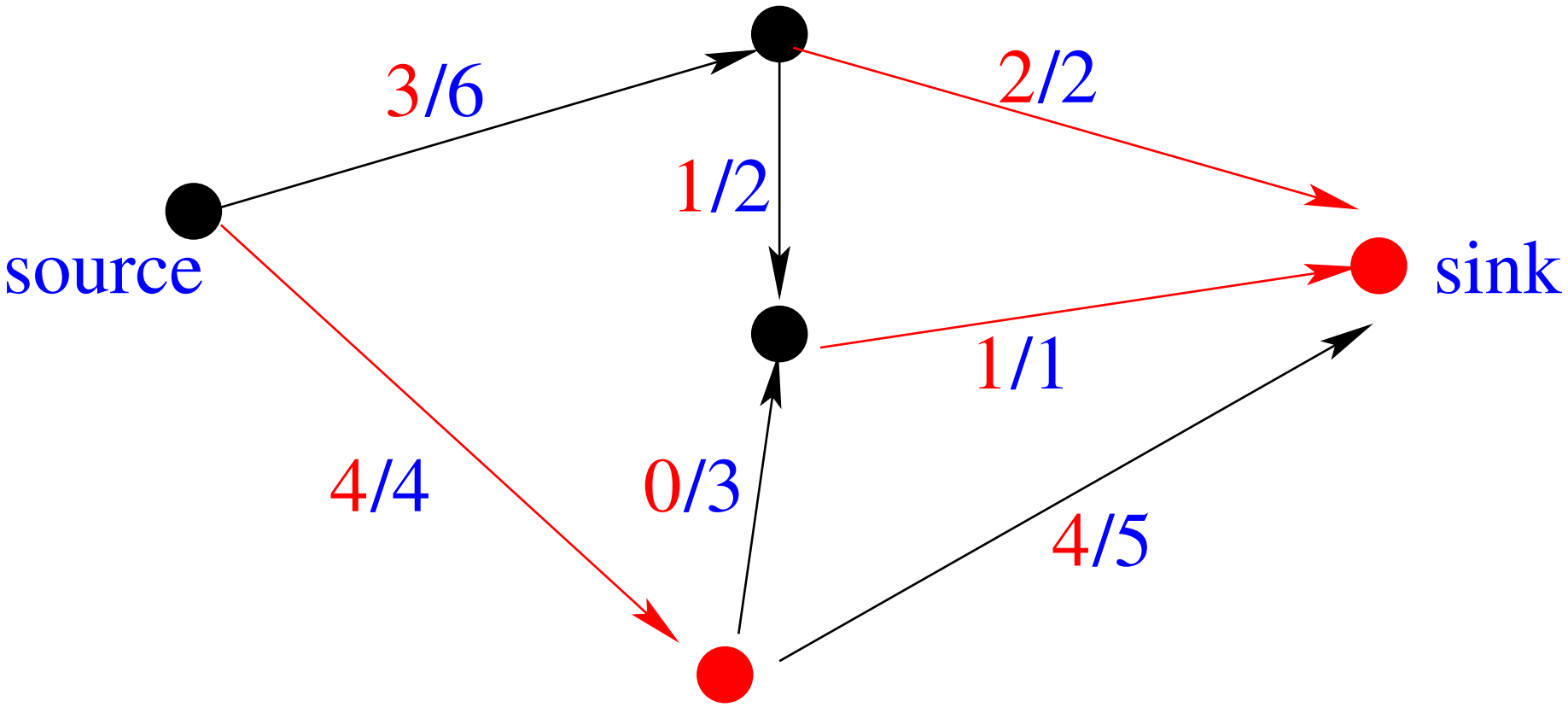


Maximum Flow / Minimum Cut



- Use of a maximum flow computation algorithm [Goldberg and Tarjan, 1988] to separate a seed of users from the remaining of the graph
- Complexity $O(n^2 m)$ (n: vertices, m: edges)

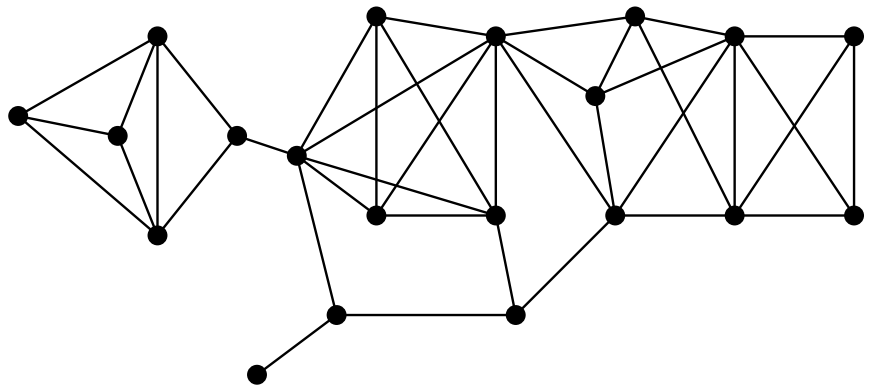
Maximum Flow / Minimum Cut



- Use of a maximum flow computation algorithm [Goldberg and Tarjan, 1988] to separate a seed of users from the remaining of the graph
- Complexity $O(n^2 m)$ (n: vertices, m: edges)

Markov Cluster Algorithm (MCL) [van Dongen, 2000]

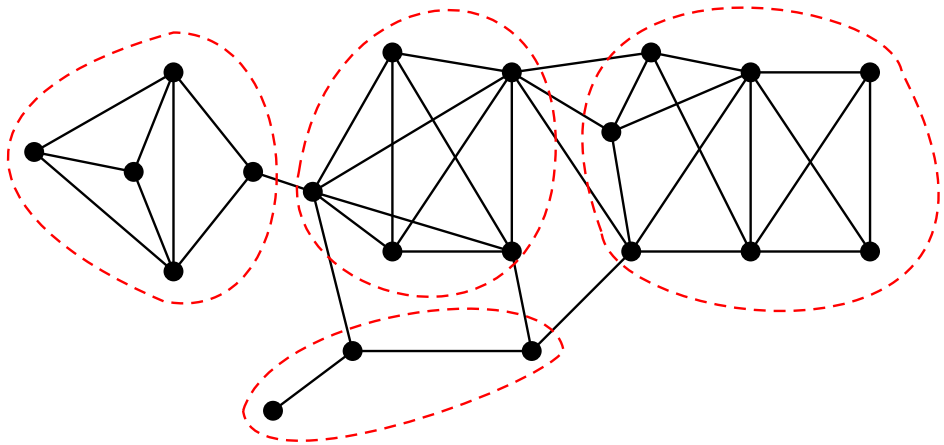
- Graph clustering algorithm
- Based as well on maximum flow simulation, in the whole graph
- Iteration of a matrix computation alternating:
  - Expansion (matrix multiplication, corresponding to flow propagation)
  - Inflation (non-linear operation to increase heterogeneity)
- Complexity: $O(n^3)$ for an exact computation, $O(n)$ for an approximate one



[van Dongen, 2000]

Markov Cluster Algorithm (MCL) [van Dongen, 2000]

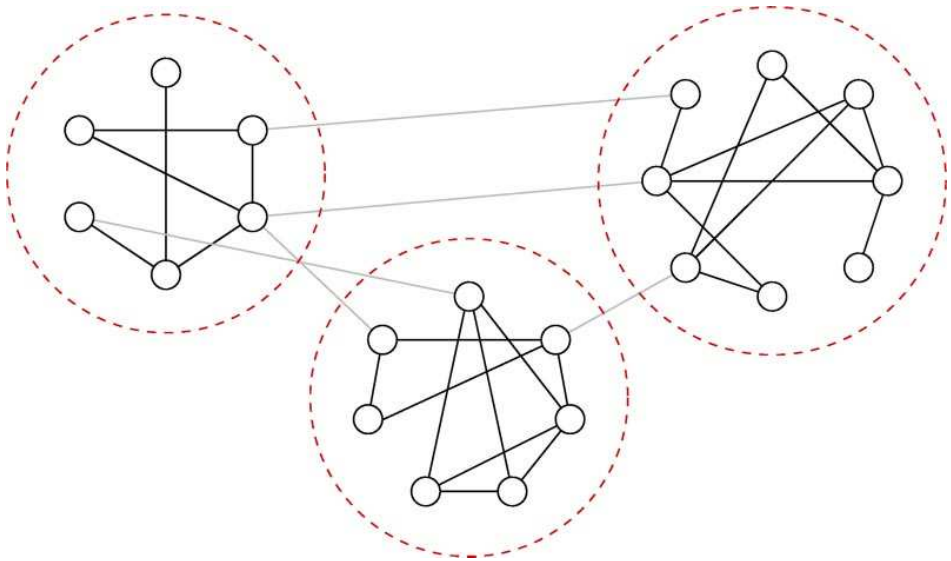
- Graph clustering algorithm
- Based as well on maximum flow simulation, in the whole graph
- Iteration of a matrix computation alternating:
  - Expansion (matrix multiplication, corresponding to flow propagation)
  - Inflation (non-linear operation to increase heterogeneity)
- Complexity: $O(n^3)$ for an exact computation, $O(n)$ for an approximate one



[van Dongen, 2000]

Deletion of the edges with the highest betwenness [Newman and Girvan, 2004]

- Top-down graph clustering algorithm
- Betwenness of an edge: number of minimal paths between two arbitrary vertices going through this edge
- General principle:
  1. Compute the betweenness of each edge in the graph
  2. Remove the edge with the highest betweenness
  3. Redo the whole process, betweenness computation included
- Complexity: $O(n^3)$ for a sparse graph



[Newman and Girvan, 2004]

Outline

1 Introduction

2 Basics of Graph Theory

3 Characteristics of Real-World Networks

4 Models of Networks

5 Graph Mining Algorithms

6 Conclusion

To remember

What you should remember

1. Most (but not all!) real-world networks:
   - are sparse
   - have small typical distance
   - have strong local clustering
2. In addition, a large class of them are scale-free
3. Three simple models of networks, modeling (and explaining?) some or all of these properties:
   - Random graphs
   - Small worlds
   - Preferential attachment
4. A collection of algorithms for mining graphs

Applications of the Models

- Epidemiology
- Network fault detection
- Efficient search in P2P networks
- ...

To go further

[Watts, 1999]: an easy-to-read book describing the small world problem and small-world models, with concrete applications

[Newman et al., 2006]: an in-detail survey of the most fundamental works on network theory, networks models, and experimentations on real-world networks

[Chakrabarti, 2003]: the particular example of the World Wide Web

Bibliographie I

L. A. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley. Classes of small-world networks. *PNAS*, 97(21):11149–11152, October 2000.

Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, October 1999.

Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1–7):107–117, 1998.

Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. Graph structure in the web. *Computer Networks*, 33(1-6):309–320, 2000.

Soumen Chakrabarti. *Mining the Web: Discovering Knowledge from Hypertext Data*. Morgan Kaufmann, San Fransisco, USA, 2003.

Bibliographie II

Peter Sheridan Dodds, Roby Muhamad, and Duncan J. Watts. An experimental study of search in global social networks. *Science*, 301 (5634):827–829, August 2003.

P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5:17–61, 1960.

M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *Proc. SIGCOMM*, pages 251–262, Cambridge, USA, August 1999.

Andrew V. Goldberg and Robert E. Tarjan. A new approach to the maximum-flow problem. *Journal of the ACM*, 35(4):921–940, October 1988.

H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabasi. The large-scale organization of metabolic networks. *Nature*, 407(6804), 2000.

Bibliographie III

Frigyes Karinthy. Chains. In *Everything is different*. 1929. Translated from Hungarian by Ádám Makkai, as reproduced in [Newman et al., 2006].

Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.

F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aaberg. The web of human sexual contacts. *Nature*, 411(6840):907–908, 2001.

Nan Lin. *Social Capital: A Theory of Social Structure and Action*. Cambridge University Press, Cambridge, United Kingdom, 2001.

M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69(2), 2004.

Mark Newman, Albert-László Barabási, and Duncan J. Watts. *The Structure and Dynamics of Networks*. Princeton University Press, 2006.

M. Smith, C. Giraud-Carrier, and B Judkins. Implicit affinity networks. In *Proc. Workshop on Information Technologies and Systems*, pages 1–7, Montreal, Canada, December 2007.

Bibliographie IV

Ray Solomonoff and Anatol Rapoport. Connectivity of random nets. *Bulletin of Mathematical Biology*, 13(2):107–117, June 1951.

Jeffrey Travers and Stanley Milgram. An experimental study of the small world problem. *Sociometry*, 34(4), December 1969.

Stijn Marinus van Dongen. *Graph Clustering by Flow Simulation*. PhD thesis, University of Utrecht, May 2000.

Duncan J. Watts. *Small Worlds*. Princeton University Press, 1999.

Duncan J. Watts and Steven H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.