Lectures 16, 17

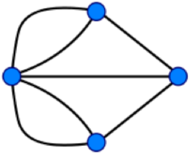
Charalampos E. Tsourakakis

3/27 & 3/29/17

Introduction to Graphs and Networks

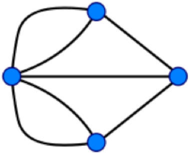
Graphs: a simple model

- entities – set of vertices
- pairwise relations among vertices
  – set of edges
- can add directions, weights,. . .
- graphs can be used to model many real datasets
  - people who are friends
  - computers that are interconnected
  - web pages that point to each other
  - proteins that interact



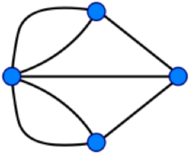
Graphs: a simple model

- entities – set of vertices
- pairwise relations among vertices
  – set of edges
- can add directions, weights,. . .
- graphs can be used to model many real datasets
  - people who are friends
  - computers that are interconnected
  - web pages that point to each other
  - proteins that interact



Graphs: a simple model

- entities – set of vertices
- pairwise relations among vertices
  – set of edges
- can add directions, weights,...
- graphs can be used to model many real datasets
  - people who are friends
  - computers that are interconnected
  - web pages that point to each other
  - proteins that interact

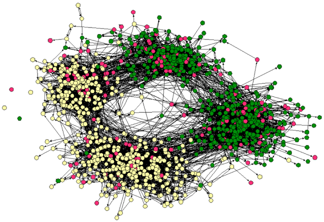


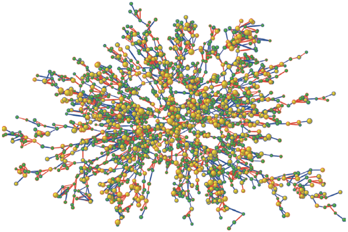
Graph theory

- graph theory started in the 18th century, with Leonhard Euler
  - the problem of Königsberg bridges
  - since then, graphs have been studied extensively

Analysis of Graph Datasets in the Past

- graphs datasets have been studied in the past
  e.g., networks of highways, social networks
  - usually these datasets were small
  - visual inspection can reveal a lot of information

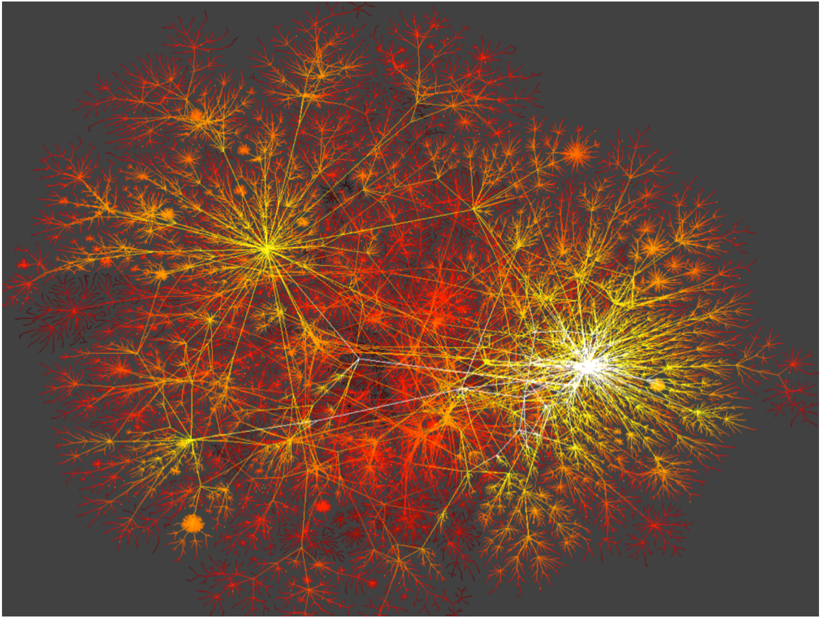




Analysis of Graph Datasets Now

- more and larger networks appear
  - products of technological advancement
    - e.g., internet, web
  - result of our ability to collect more, better-quality, and more complex data
    - e.g., gene regulatory networks
- networks of thousands, millions, or billions of nodes
  - impossible to visualize

The Internet map

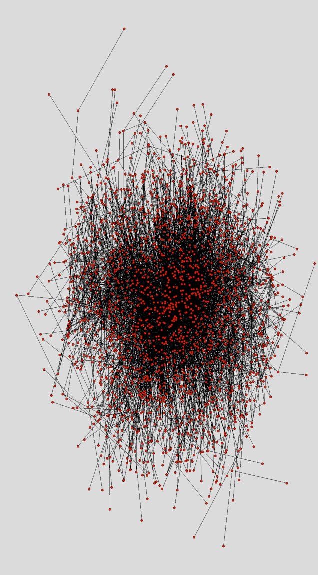


Types of Networks

- social networks
- knowledge and information networks
- technology networks
- biological networks

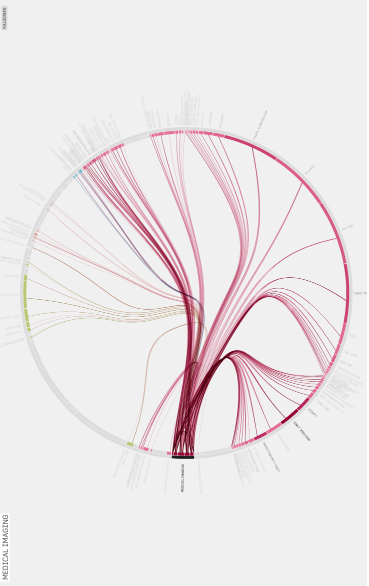
Social Networks

- links denote a social interaction
  - networks of acquaintances
  - collaboration networks
    - actor networks
    - co-authorship networks
    - director networks
  - phone-call networks
  - e-mail networks
  - IM networks
  - sexual networks



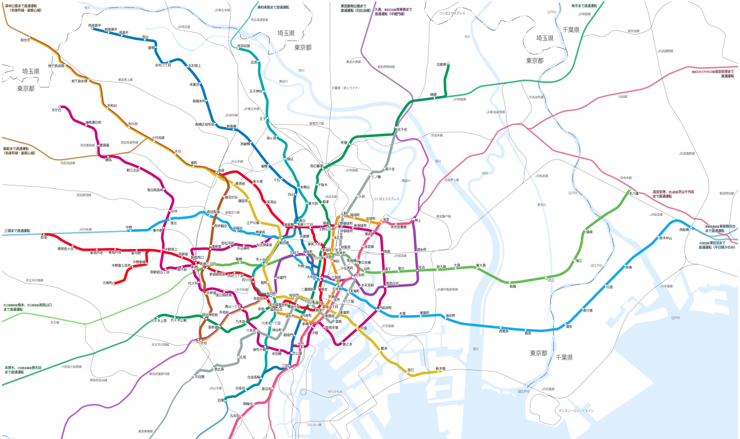
Knowledge and Information Networks

- nodes store information, links associate information
  - citation network (directed acyclic)
  - the web (directed)
  - peer-to-peer networks
  - word networks
  - networks of trust
  - software graphs
  - bluetooth networks
  - home page/blog networks

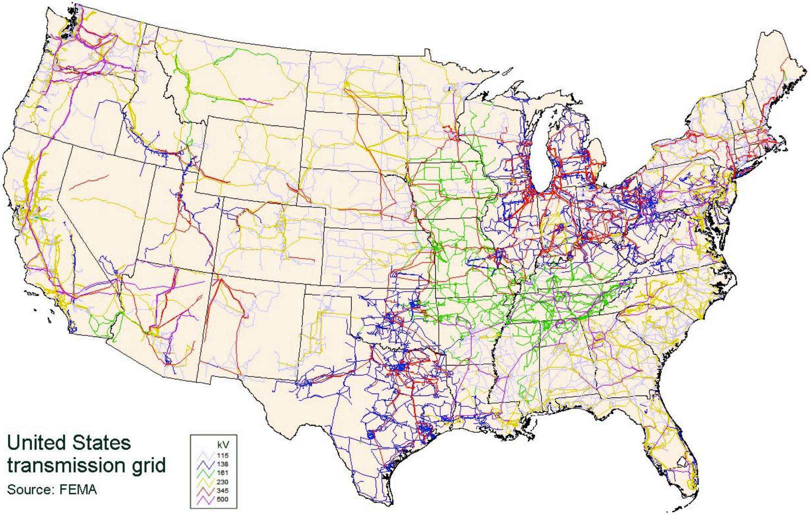


Technological Networks

- networks built for distribution of a commodity
    - the internet, power grids, telephone networks
    - airline networks, transportation networks

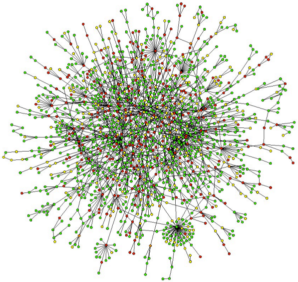


US power grid



Biological Networks

- biological systems represented as networks
  - protein-protein interaction networks
  - gene regulation networks
  - gene co-expression networks
  - metabolic pathways
  - the food web
  - neural networks



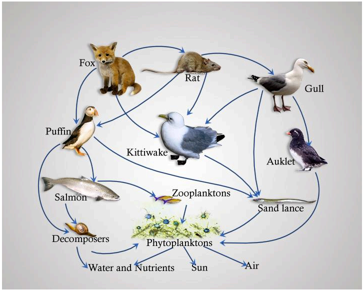
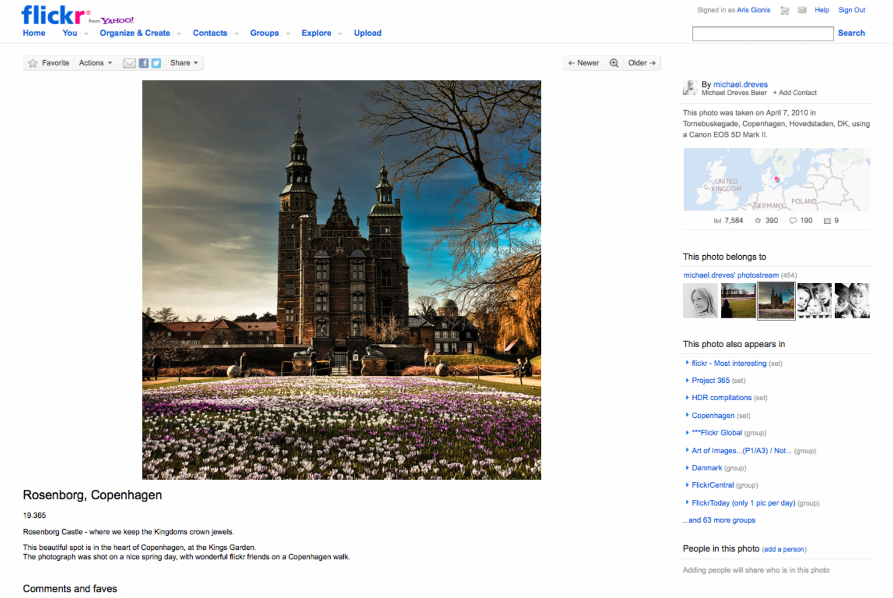


Photo-sharing Site



What is the Underlying Graph?

- nodes: photos, tags, users, groups, albums, sets, collections, geo, query, ...
- edges: upload, belong, tag, create, join, contact, friend, family, comment, fave, search, click, ...
- also many interesting induced graphs
  - tag graph: based on photos
  - tag graph: based on users
  - user graph: based on favorites
  - user graph: based on groups
- which graph to pick — not an easy choice

Recurring Theme

- social media, user-generated content
- user interaction is composed by many atomic actions
  - post, comment, like, mark, join, comment, fave, thumps-up, . . .
  - generates all kind of interesting graphs to mine

Network Science

- the world is full with networks
- what do we do with them?
  - understand their topology and measure their properties
  - study their evolution and dynamics
  - create realistic models
  - create algorithms that make use of the network structure

Outline

- Random graphs as models of real-world networks Babis
  - Properties of real-world networks
  - Erdös-Rényi graphs
  - Models of real world networks
  - Applications
- Algorithm design for large-scale networks Aris
  - Graph sparsifiers
  - Graph partitioning
  - Dense subgraphs
  - Applications

Properties of Real-World Networks

Properties of Real-World Networks

Diverse collections of graphs arising from different phenomena

Are there any typical patterns?

- Static networks
  1. Heavy tails
  2. Clustering coefficients
  3. Communities
  4. Small diameters
- Time-evolving networks
  1. Densification
  2. Shrinking diameters
- Web graph
  1. Bow tie structure
  2. Bipartite cliques

Heavy tails

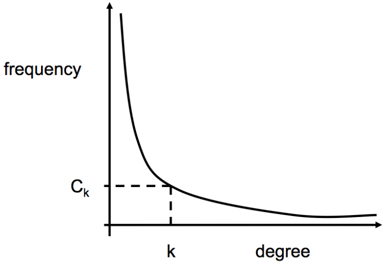
- *What do the proteins in our bodies, the Internet, a cool collection of atoms and sexual networks have in common? One man thinks he has the answer and it is going to transform the way view the world.*

Scientist 2002

Albert-László Barabási

Degree distribution

- C_k = number of vertices with degree k



- problem : find the probability distribution that fits best the observed data

Degree distribution

- C_k = number of vertices with degree k, then

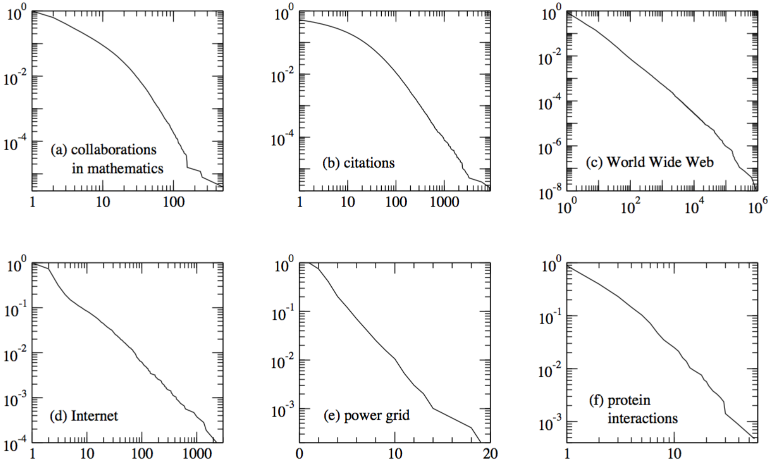
$$C_k = ck^{-\gamma}$$

with $\gamma > 1$, or

$$\ln C_k = \ln c - \gamma \ln k$$

- so, plotting $\ln C_k$ versus $\ln k$ gives a straight line with slope $-\gamma$
- heavy-tail distribution : there is a non-negligible fraction of nodes that has very high degree (hubs)
- scale free : average is not informative

Power-law Degree Distribution



Power-laws appear in a wide variety of networks (Source [Newman, 2003]). These degree distributions stand in sheer contrast to Erdös-Rényi random graphs.

Power-law Degree Distribution

Do the degrees follow a power-law distribution?

There were three major problems with the initial studies that indicated that degrees follow a power law.

- The way in which the graphs were generated in the traceroute studies produces power-law distributions, even for regular graphs [Lakhina et al., 2003].
- Even if the true distribution is a power law, the methodology used to determine the exponent of the power law distribution. See [Clauset et al., 2009] for a proper methodology.
- Other distributions could potentially fit the data better but were not considered,.e.g, lognormal, double Pareto lognormal.

Disclaimer: We will be referring to these distributions as heavy-tailed, avoiding a specific characterization of the distribution.

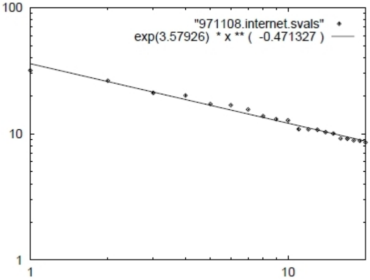
Maximum Degree

- for random graphs, the maximum degree is highly concentrated around the average degree z
- for power-law graphs

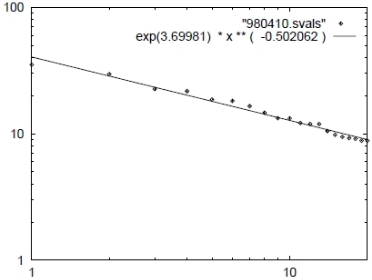
$$d_{\max} \approx n^{1/(\alpha-1)}$$

- hand-waving argument: solve $n \Pr[X \geq d] = 1$

Heavy Tails, Eigenvalues



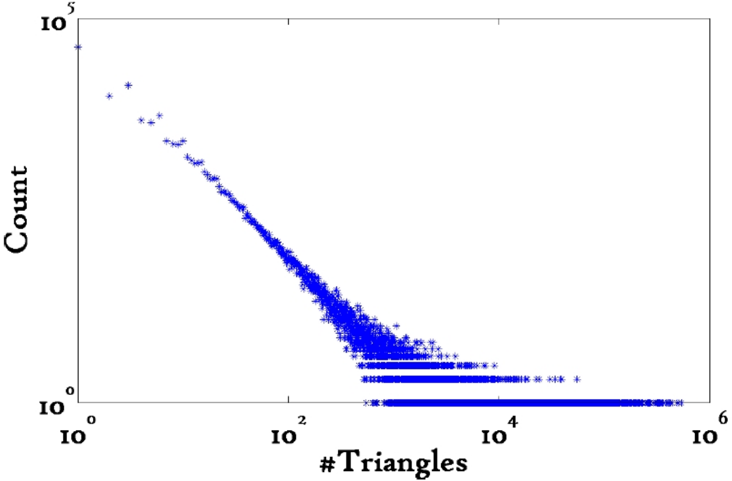
(a) Int-11-97



(b) Int-04-98

Eigenvalues of the Internet graph follow a power law
[M. Faloutsos, 1999]

Heavy Tails, Triangles



Triangle distribution of Flickr. Again, heavy tails emerge [Tsourakakis, 2008].

Community Structure

- intuitively a subset of vertices that are more connected to each other than to other vertices in the graph
- a proposed measure is clustering coefficient

$$C_1 = \frac{3 \times \text{number of triangles in the network}}{\text{number of connected triples of vertices}}$$

- captures "transitivity of clustering"
- if u is connected to v and
  v is connected to w, it is also likely that
  u is connected to w
- community structure at local level

Community Structure

- alternative definition
- local clustering coefficient

$$C_i = \frac{\text{number of triangles connected to vertex } i}{\text{number of triples centered at vertex } i}$$

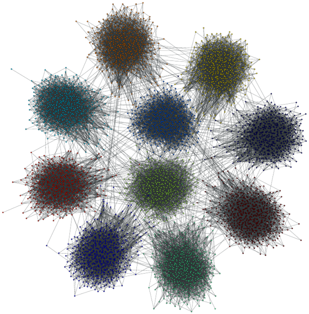
- global clustering coefficient

$$C_2 = \frac{1}{n} \sum_i C_i$$

- community structure is captured by large values of clustering coefficient

Community Structure

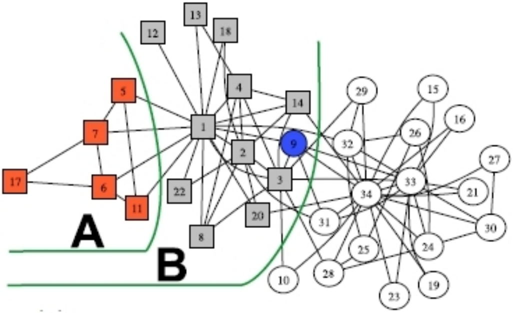
Loose definition: A community is a set of vertices S which is "densely" interconnected and "sparsely" connected to the rest of the vertices.



Artificial communities (Source `http://projects.skewed.de/graph-tool/`)

Community Structure

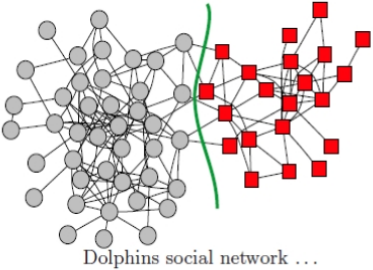
Leskovec, Lang, Dasgupta and Mahoney [Leskovec et al., 2009] study the community structure of an extensive collection of real-world networks.

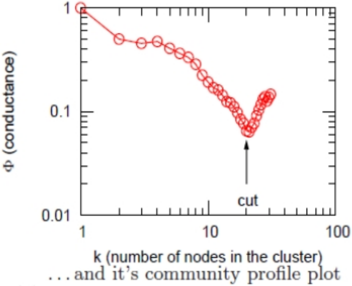


Zachary network, community structure (Source [Leskovec et al., 2009]).

Community Structure

[Leskovec et al., 2009] introduce the *network community profile plot*. It characterizes the "best" possible community over a wide range of scales.

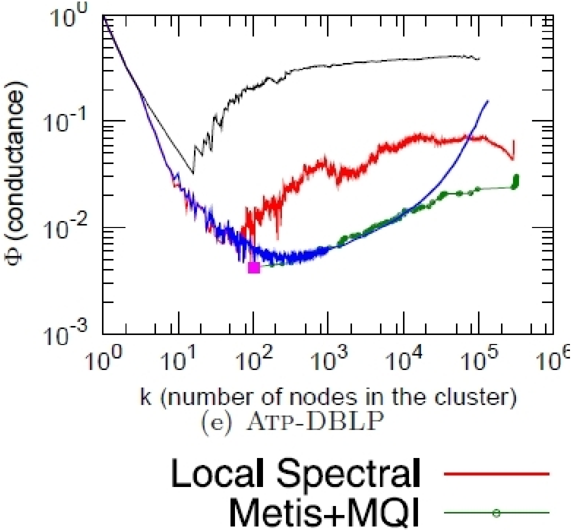




Dolphins network and its NCP (Source [Leskovec et al., 2009]).

Community Structure

- Do large-scale real-world networks have this nice structure? NO!



(e) ATP-DBLP

Community Structure

Important findings of [Leskovec et al., 2009].

1. Up to a certain size k ($k \sim 100$ vertices) there exist good cuts and as the size increases so does the quality of the community
2. At the size k we observe the best possible community. These communities are typically connected to the remainder with a single edge.
3. Above the size k, the community quality decreases. This is because they blend in and gradually disappear.

small-world phenomena

small worlds : graphs with short paths

- Stanley Milgram (1933-1984)
  "The man who shocked the world"
- obedience to authority (1963)
- small-world experiment (1967)

small-world experiments

- letters were handed out to people in Nebraska to be sent to a target in Boston
- people were instructed to pass on the letters to someone they knew on first-name basis
- the letters that reached the destination (64 / 296) followed paths of length around 6
- six degrees of separation : (play of John Guare)
- also:
  - the Kevin Bacon game
  - the Erdös number
- small-world project: http://smallworld.columbia.edu/index.html

small diameter

proposed measures

- diameter : largest shortest-path over all pairs
- effective diameter : upper bound of the shortest path of 90% of the pairs of vertices
- average shortest path : average of the shortest paths over all pairs of vertices
- characteristic path length : median of the shortest paths over all pairs of vertices
- hop-plots : plot of $|N_h(u)|$, the number of neighbors of u at distance at most h, as a function of h
  [M. Faloutsos, 1999] conjectured that it grows exponentially and considered hop exponent

measurements on real graphs

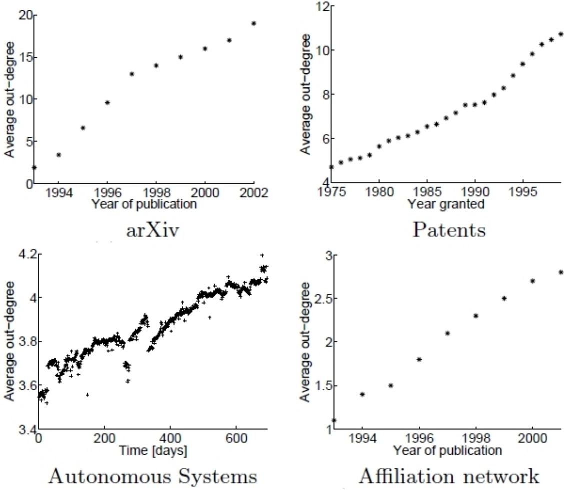
	n	m	γ	C_1	C_2	ℓ
film actors	449 913	25 516 482	2.3	0.20	0.78	3.48
internet	10 697	31 992	2.5	0.03	0.39	3.31
protein interactions	2 115	2 240	2.4	0.07	0.07	6.80

[Newman, 2003]

other properties

- degree correlations
- distribution of size of connected components
- resilience
- eigenvalues
- distribution of motifs

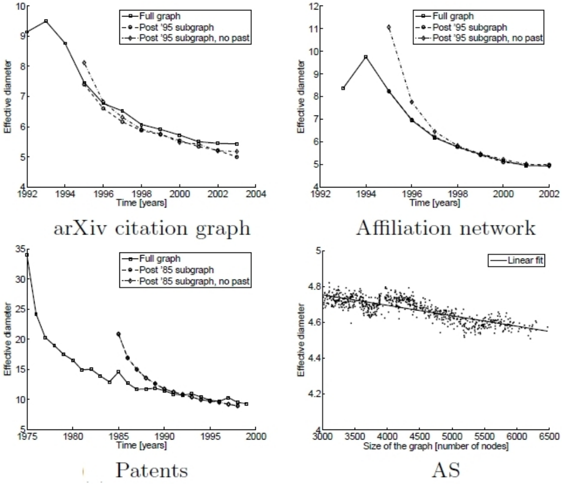
Densification



densification power law (Source [Leskovec et al., 2005b])

$$|E_t| \propto |V_t|^{\alpha} \qquad 1 \leq \alpha \leq 2$$

Shrinking diameters

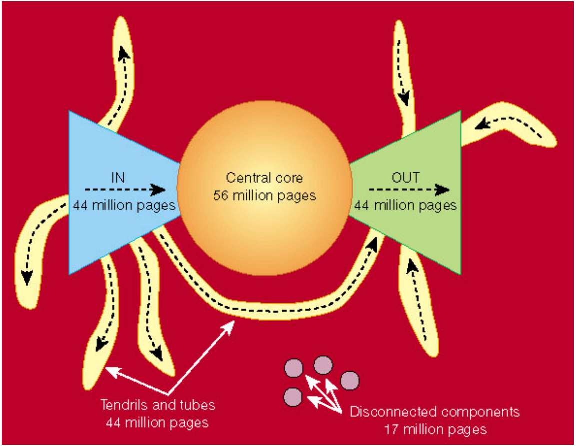


diameter is shrinking (Source [Leskovec et al., 2005b])

Web Graph

- The Web graph is a particularly important real-world network. *Few events in the history of computing have wrought as profound an influence on society as the advent and growth of the World Wide Web* [Kleinberg et al., 1999a].
  - Vertices correspond to static Web pages
  - Directed edge (i, j) corresponds to a link from page i to page j.
- We will discuss two structural properties of the Web graph:
  1. The bow-tie structure [Broder et al., 2000]
  2. Abundance of bipartite cliques [].

Web is a bow-tie

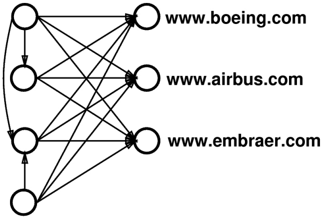


(Source [nat, 2000, Broder et al., 2000])

Bipartite subgraphs

- Websites that are part of the same community frequently do not reference one another (competitive reasons, disagreements, ignorance) [Kumar et al., 1999].
- Similar Websites are *co-cited*.

Therefore, Web communities are characterized by dense directed bipartite subgraphs.



(Source [Kleinberg et al., 1999a])

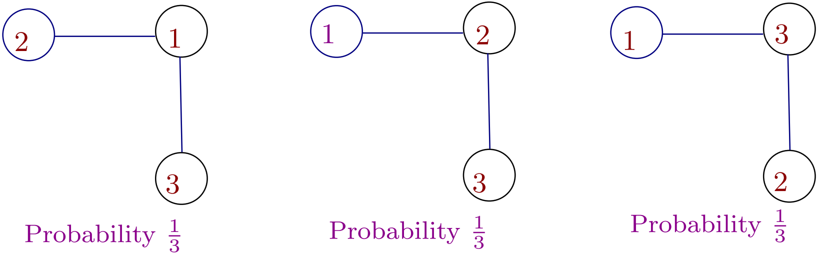
random graphs

random graphs

random graphs

A random graph is a set of graphs together with a probability distribution on that set.

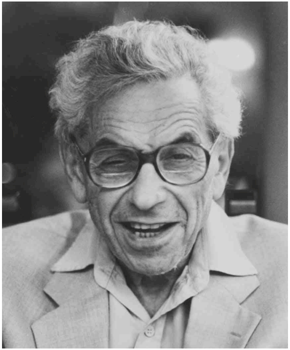
Example:



A random graph on $\{1, 2, 3\}$ with 2 edges with the uniform distribution.

random graphs

- Erdős-Rényi random graph model



Paul Erdős
1913 – 1996

Alfréd Rényi
1921 – 1970

random graphs

- the $G(n, p)$ model:
- n : the number of vertices
- $0 \leq p \leq 1$: probability
- for each pair (u, v), independently generate the edge (u, v) with probability p
- $G(n, p)$ a family of graphs, in which a graph with m edges appears with probability $p^m(1-p)^{\binom{n}{2}-m}$
- the $G(n, m)$ model: related, but not identical

random graphs

- the $G(n, p)$ model:
- n : the number of vertices
- $0 \leq p \leq 1$: probability
- for each pair (u, v), independently generate the edge (u, v) with probability p
- $G(n, p)$ a family of graphs, in which a graph with m edges appears with probability $p^m(1-p)^{\binom{n}{2}-m}$
- the $G(n, m)$ model: related, but not identical

random graphs

- the $G(n, p)$ model:
- n : the number of vertices
- $0 \leq p \leq 1$: probability
- for each pair (u, v), independently generate the edge (u, v) with probability p
- $G(n, p)$ a family of graphs, in which a graph with m edges appears with probability $p^m(1-p)^{\binom{n}{2}-m}$
- the $G(n, m)$ model: related, but not identical

properties of random graphs

- a property P holds almost surely (for almost every graph) if

$$\lim_{n\to\infty} \Pr[G \text{ has } P] = 1$$

- which properties hold as p increases?
- threshold phenomena : many properties appear suddenly
- there exist a probability p_c such that

  for $p < p_c$ the property does not hold a.s.

  for $p > p_c$ the property holds a.s.

properties of random graphs

- a property P holds almost surely (for almost every graph) if

$$\lim_{n\to\infty} \Pr[G \text{ has } P] = 1$$

- which properties hold as p increases?
- threshold phenomena : many properties appear suddenly
- there exist a probability p_c such that

  for $p < p_c$ the property does not hold a.s.

  for $p > p_c$ the property holds a.s.

the giant component

- let $z = np$ be the average degree
- if $z < 1$ the largest component has size $O(\log n)$ a.s.
- if $z > 1$ the largest component has size $\Theta(n)$ a.s.;
  the second largest component has size $O(\log n)$ a.s.
- if $z = \omega(\log n)$ the graph is connected a.s.

phase transition

- if $z = 1$ there is a phase transition
  - the largest component has size $O(n^{2/3})$
  - the sizes of the components follow a power-law

Phase Transition, proof sketch

Michael Krivelevich Benny Sudakov

The phase transition in random graphs - a simple proof

The Erdös-Rényi paper, which launched the modern theory of random graphs, has had enormous influence on the development of the field and is generally considered to be a single most important paper in Probabilistic Combinatorics, if not in all of Combinatorics.

Phase Transition, proof sketch

Krivelevich and Sudakov [Krivelevich and Sudakov, 2013] give a simple proof for the transition based on running the depth first search (DFS) algorithm on G.

- S : vertices whose exploration is complete
- T : unvisited vertices
- $U = V - (S \cup T)$: vertices in stack

Observation: The set U always spans a path, since when a vertex u is added in U, it happens because u is a neighbor of the last vertex v in U; thus, u augments the path spanned by U, of which v is the last vertex.
Epoch is the period of time between two consecutive emptyings of U. Each epoch corresponds to a connected component.

Phase Transition, proof sketch

Lemma

Let $\epsilon > 0$ be a small enough constant. Consider the sequence $\bar{X} = (X_i)_{i=1}^{N}$ of i.i.d. Bernoulli random variables with parameter p.

1. *Let $p = \frac{1-\epsilon}{n}$. Let $k = \frac{7}{\epsilon^2} \ln n$. Then* **whp** *there is no interval of length kn in $[N]$, in which at least k of the random variables X_i take value 1.*
2. *Let $p = \frac{1+\epsilon}{n}$. Let $N_0 = \frac{\epsilon n^2}{2}$. Then* **whp** $\left| \sum_{i=1}^{N_0} X_i - \frac{\epsilon(1+\epsilon)n}{2} \right| \leq n^{2/3}$.

Phase Transition, proof sketch

We will use the following simple but extremely useful tools.

Lemma (Union Bound)

For any events $A_1, \ldots, A_n$, $\mathbf{Pr}\left[A_1 \cup \ldots A_n\right] \leq \sum_{i=1}^{n} \mathbf{Pr}\left[A_i\right]$.

Lemma (Chebyshev's Inequality)

Let X *be a random variable with finite expectation* $\mathbb{E}\left[X\right]$ *and finite non-zero variance* $\mathbb{V}ar\left[X\right]$. *Then for any* $t > 0$,

$$\mathbf{Pr}\left[\left|X - \mathbb{E}\left[X\right]\right| \geq t\right] \leq \frac{\mathbb{V}ar\left[X\right]}{t^2}.$$

Lemma (Chernoff bound, upper tail)

Let $0 \leq \epsilon \leq 1$. *Then,*

$$\mathbf{Pr}\left[Bin(n,p) \geq (1+\epsilon)np\right] \leq e^{-\frac{\epsilon^2}{3}np}$$

Phase Transition, proof sketch

Proof.

- Fix an interval I of length kn in $[N]$, $N = \binom{n}{2}$. Then, $\sum_{i \in I} X_i \sim Bin(kn, p)$.
  1. Apply the Chernoff bound to the upper tail of $B(kn, p)$.
  2. Apply the union bound over all $(N - k + 1)$ possible intervals of length kn to upper bound the probability of the existence of an interval violating the assertion of the lemma.

  $$(N - k + 1)Pr[B(kn, p) \geq k] < n^2 \cdot e^{-\frac{\epsilon^2}{3}(1-\epsilon)k} = o(1)\,,$$

  for small enough $\epsilon > 0$.
- The sum $\sum_{i=1}^{N_0} X_i$ is distributed binomially with parameters N_0 and p. Hence, its expectation is $N_0 p = \frac{\epsilon n^2 p}{2} = \frac{\epsilon(1+\epsilon)n}{2}$, and its standard deviation is of order n. Applying the Chebyshev inequality, we get the required estimate.

Phase Transition, proof sketch

Proof.

CASE I: $p = \frac{1-\epsilon}{n}$

- Assume to the contrary that G contains a connected component C with more than $k = \frac{7}{\epsilon^2} \ln n$ vertices.
- Consider the moment inside this epoch when the algorithm has found the $(k+1)$-st vertex of C and is about to move it to U.
- Denote $\Delta S = S \cap C$ at that moment. Then $|\Delta S \cup U| = k$, and thus the algorithm got exactly k positive answers to its queries to random variables X_i during the epoch, with each positive answer being responsible for revealing a new vertex of C, after the first vertex of C was put into U in the beginning of the epoch.

$\square$

Phase Transition, proof sketch

Proof.

- At that moment during the epoch only pairs of edges touching $\Delta S \cup U$ have been queried, and the number of such pairs is therefore at most $\binom{k}{2} + k(n-k) < kn$. It thus follows that the sequence $\bar{X}$ contains an interval of length at most kn with at least k 1's inside – a contradiction to Property 1 of Lemma 1.

<u>CASE II: $p = \frac{1+\epsilon}{n}$</u>

- Same idea. Assume the result does not hold and reach a contradiction by examining carefully the number of queries.

$\square$

Degree Distribution

- degree distribution : binomial

$$C_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

- The limit distribution of the normalized binomial distribution $Bin(n, p)$ is the normal distribution provided that $np(1-p) \to +\infty$ as $n \to +\infty$.
- If $p = \frac{\lambda}{n}$ the limit distribution of $Bin(n, p)$ is the Poisson distribution.

Degree Distribution

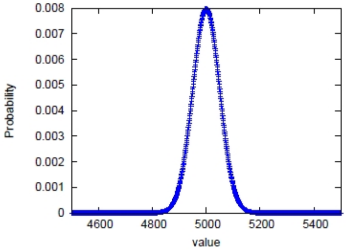


FIGURE 1. The Binomial distribution $B(10000, 0.5)$.

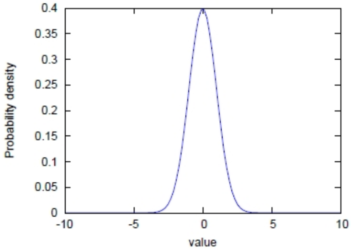


FIGURE 2. The Standard normal distribution $N(0, 1)$.

$Bin(10000, 0.5)$ and Gaussian(0,1)

Degree Distribution

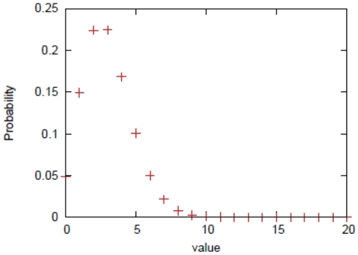


FIGURE 3. The Binomial distribution $B(1000, 0.003)$.

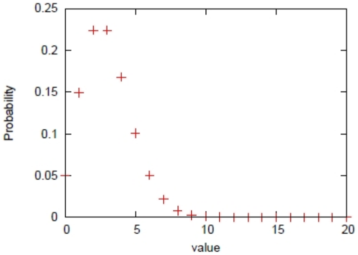


FIGURE 4. The Poisson distribution $P(3)$.

$Bin(1000, 0.003)$ and Poisson(3)

Degree Distribution

Theorem

Let $p = \frac{\log n}{n} \times \omega(n)$ *where* $\omega(n) \to +\infty$ *arbitrarily slowly. Let* $x \in G$ *be fixed. Fix* $\epsilon > 0$*. Then in* $G(n,p)$ **whp** *for all vertices* x

$$deg(x) \sim (n-1)p.$$

Theorem (McKay, Wormald [McKay and Wormald, 1997])

Let X_k *be the number of vertices of degree* k *in* $G(n,p)$ *when* $p = \frac{c}{n}$*, with* $c > 0$ *constant. Then* **whp** *for* $k = 0, 1, \ldots$

$$\frac{c^k e^{-c}}{k!} \leq \frac{X_k}{n} \leq (1+\epsilon)\frac{c^k e^{-c}}{k!},$$

as $n \to +\infty$.

random graphs and real datasets

- a beautiful and elegant theory studied exhaustively
- have been used as idealized generative models
- unfortunately, they don't capture reality...

Models

1. Classics
   - Grown versus static random graphs
   - Growth with Preferential attachment
   - Structure+Randomness $\rightarrow$ Small world networks
2. More models
   - Cooper-Frieze model
   - Copying model
   - Chung-Lu model
   - Kronecker graphs
   - Affiliation networks
   - Forest-fire model
   - Random Apollonian Networks (planar)

CHNKS Model

Callaway, Hopcroft, Kleinberg, Newman and Strogatz [Callaway et al., 2001] propose a simple growth model for a random graph without preferential attachment.

Main thesis of [Callaway et al., 2001]: Grown graphs, however randomly they are constructed, are fundamentally different from their static random graph counterparts.

CHNKS Model

- Start with 0 vertices at time 0.
- At time t, a new vertex is created and with probability δ we add a random edge by choosing two existing vertices uniformly at random.

CHNKS Model

Let $d_k(t)$ be the number of vertices of degree k at time t. Then,

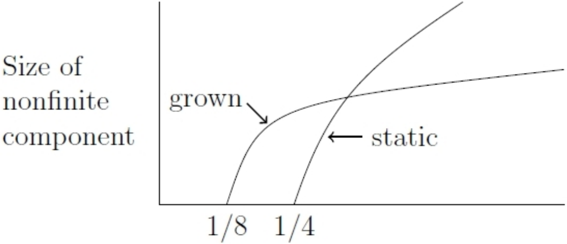
$$\mathbb{E}\left[d_0(t+1)\right] = \mathbb{E}\left[d_0(t)\right] + 1 - \delta \frac{2\mathbb{E}\left[d_0(t)\right]}{t},$$

$$\mathbb{E}\left[d_k(t+1)\right] = \mathbb{E}\left[d_k(t)\right] + \delta \Big(\frac{2\mathbb{E}\left[d_{k-1}(t)\right]}{t} - \frac{2\mathbb{E}\left[d_k(t)\right]}{t} \Big).$$

Turns out that

$$\frac{\mathbb{E}\left[d_k(t)\right]}{t} = \frac{1}{2\delta + 1} \Big(\frac{2\delta}{2\delta + 1} \Big)^k.$$

CHNKS Model



Size of giant component for a CHNKS random graph and a static random graph with the same degree distribution

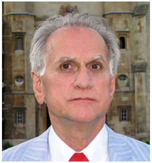
Why are grown and static random graphs so different?

Intuition: *There is a positive correlation between the degrees of connected vertices in the grown grapholder vertices tend to have higher degree, and to link with other highdegree vertices, merely by virtue of their age*

Preferential Attachment

R. Albert

L. Barabási



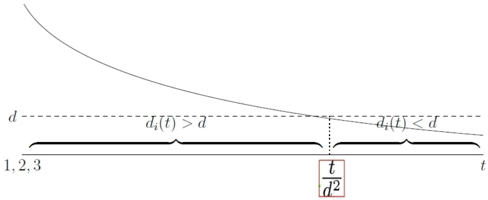
B. Bollobás

O. Riordan

Growth model: At time n, vertex n is added to the graph, *one* edge is attached to the new vertex and to a vertex selected at random with probability proportional to its degree. We obtain a sequence of graphs $\{G_1^{(n)}\}$.

The case of $G_m^{(n)}$ where instead of a single edge we add m edges reduces to $G_1^{(n)}$ by creating a $G_1^{(nm)}$ and then collapsing vertices $km, km-1, \ldots, (k-1)m+1$ to create vertex k.

Preferential Attachment



At time t, vertices numbered 1 to $\frac{1}{d^2}$ have degrees greater than d

Heuristic Analysis: Let $deg_i(t)$ be the *expected* degree of the i-th vertex at time t. The probability an edge is connected to i is $\frac{deg_i(t)}{2t}$. Therefore,

$$\frac{\partial deg_i(t)}{\partial t} = \frac{deg_i(t)}{2t}.$$

The solution is $deg_i(t) = \sqrt{\frac{t}{i}}$.

Preferential Attachment

$$\int_0^d \mathbf{Pr}\left[degree = d\right]\partial d = \mathbf{Pr}\left[degree < d\right] = 1 - \frac{1}{d^2},$$

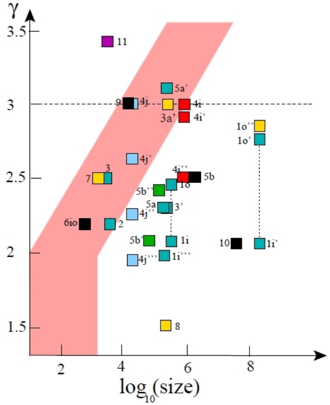
and by taking the derivative

$$\mathbf{Pr}\left[degree = d\right] = \frac{\partial}{\partial d}\left(1 - \frac{1}{d^2}\right) = \frac{2}{d^3}.$$

Power law distribution!

These results can be proved rigorously using the Linearized Chord Diagrams (LCD) model and also prove strong concentration around the expectation using martingales.

Generalized Preferential Attachment



Log-linear plot of the exponents of all the networks reported as having power-law (Source [Dorogovtsev and Mendes, 2002])

Many real world networks have a power-law slope $2 < \alpha < 3$

Generalized Preferential Attachment

How can we tune the power-law slope?

Buckley and Osthus [Buckley and Osthus, 2004] analyze a modified preferential attachment process where $\alpha > 0$ is a *fitness* parameter. When t vertex comes in, it chooses i according to

$$\mathbf{Pr}\left[t \text{ chooses } i\right] = \begin{cases} \frac{deg_{t-1}(i)+\alpha-1}{(\alpha+1)t-1}, & \text{if } 1 \le i \le t-1 \\ \frac{\alpha}{(\alpha+1)t-1}, & \text{if } i = t \end{cases}.$$

- For $\alpha = 1$ we obtain the Barabási-Albert/Bollobás-Riordan $G_1^{(n)}$model.
- The power law slope is $2 + \alpha$.

Generalized Preferential Attachment

Furthermore, the clustering coefficient of $G_m^{(n)}$ in expectation is $\frac{(m-1)\log^2 n}{8n}$ and therefore tends to 0 [Bollobás and Riordan, 2003].
This can also be fixed by generalizing the model, see [Holme and Kim, 2002, Ostroumova et al., 2012].

- Triangle formation: If an edge between v and u was added in the previous preferential attachment step, then add one more edge from v to a randomly chosen neighbor of u.

Holme-Kim Model

• Perform a preferential attachment step
• Then perform with probability β_t another preferential attachment step or a triangle formation step with probability $1 - \beta_t$.

Diameter for PA and GPA is $\frac{\log n}{\log \log n}$ and $\log n$ respectively.

Generalized Preferential Attachment

Furthermore, the clustering coefficient of $G_m^{(n)}$ in expectation is $\frac{(m-1)\log^2 n}{8n}$ and therefore tends to 0 [Bollobás and Riordan, 2003].
This can also be fixed by generalizing the model, see [Holme and Kim, 2002, Ostroumova et al., 2012].

- Triangle formation: If an edge between v and u was added in the previous preferential attachment step, then add one more edge from v to a randomly chosen neighbor of u.

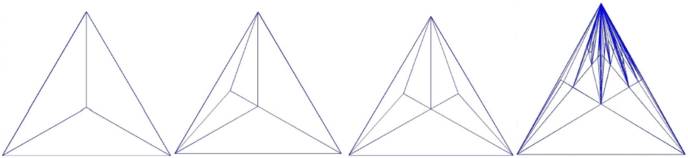
Holme-Kim Model

- Perform a preferential attachment step
- Then perform with probability β_t another preferential attachment step or a triangle formation step with probability $1 - \beta_t$.

Diameter for PA and GPA is $\frac{\log n}{\log \log n}$ and $\log n$ respectively.

Random Apollonian Networks

Are there power-law planar graphs?



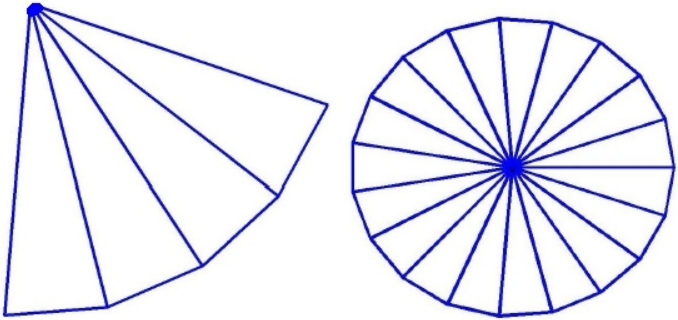
(a) $t = 1$ (b) $t = 2$ (c) $t = 3$ (d) $t = 100$

Snapshots of a Random Apollonian Network (RAN) at:
(a) t = 1 (b) t = 2 (c) t = 3 (d) t = 100.

- At time $t + 1$ we choose a face F uniformly at random among the faces of G_t. Let (i, j, k) be the vertices of F. We add a new vertex inside F and we connect it to i, j, k.

Random Apollonian Networks

Preferential attachment mechanism



What each vertex “sees” (boundary and the rest respectively)

Random Apollonian Networks

Theorem (A. Frieze, CET)

Let $Z_k(t)$ denote the number of vertices of degree k at time t, $k \geq 3$. For any $t \geq 1$ and any $k \geq 3$ there exists a constant b_k depending on k such that

$$|\mathbb{E}\left[Z_k(t)\right] - b_k t| \leq K, \quad \textit{where} \quad K = 3.6.$$

Furthermore, for t sufficiently large and any $\lambda > 0$

$$\mathbf{Pr}\left[|Z_k(t) - \mathbb{E}\left[Z_k(t)\right]| \geq \lambda\right] \leq e^{-\frac{\lambda^2}{72t}}. \tag{1}$$

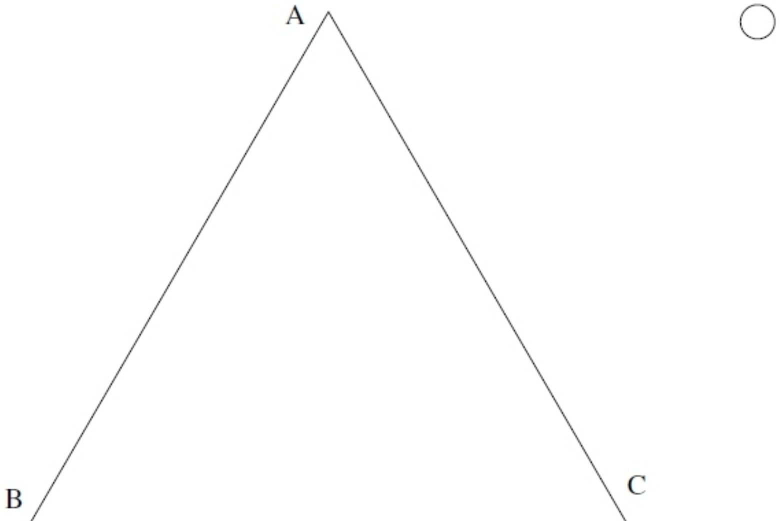
Corollary (A. Frieze, CET)

The diameter $d(G_t)$ of G_t satisfies asymptotically **whp**

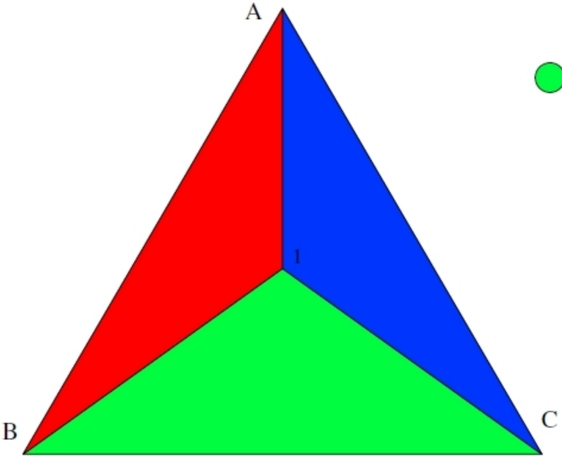
$$\mathbf{Pr}\left[d(G_t) > 7.1 \log t\right] \to 0.$$

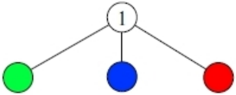
Random Apollonian Networks

The key idea is to establish a bijection with random ternary trees.

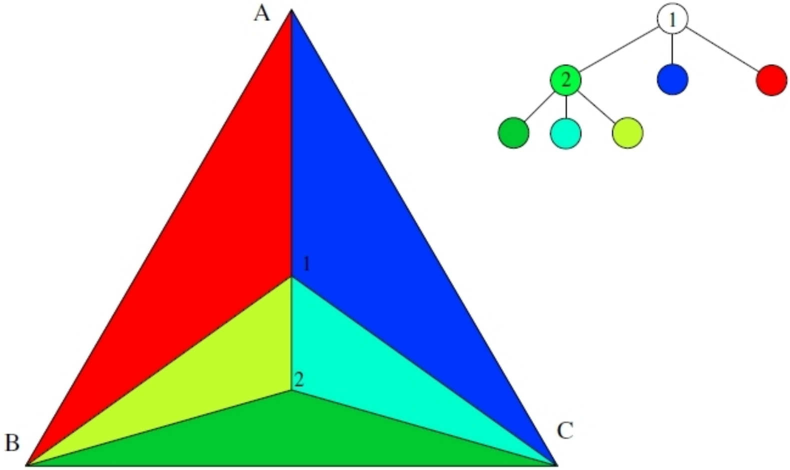


Random Apollonian Networks

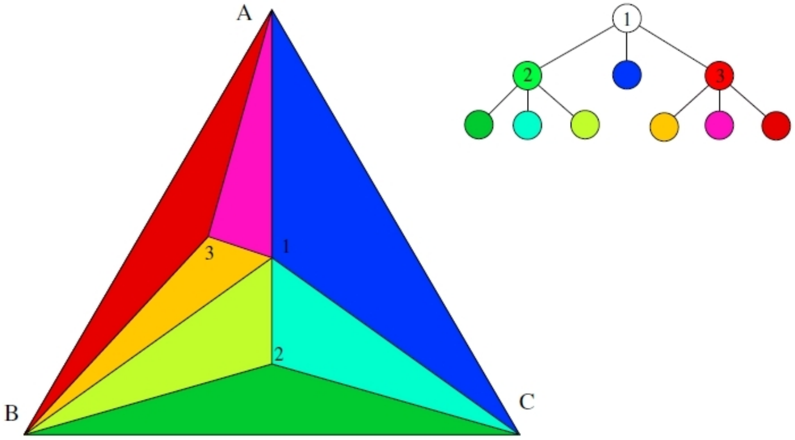




Random Apollonian Networks



Random Apollonian Networks



Small World Models

Duncan Watts

Steven Strogatz

Construct a network with

- small diameter
- positive density of triangles

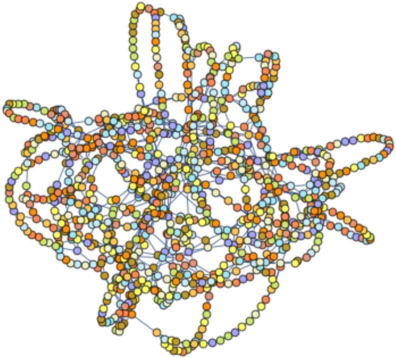
Small World Models

Why shoud we want to construct a network with

- small diameter,
- positive density of triangles?

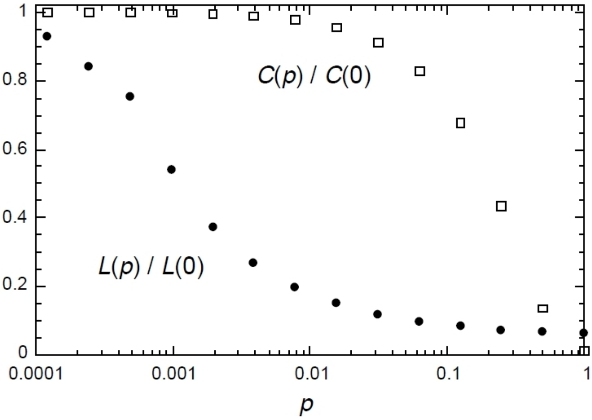
Graph	$\sim \lvert V \rvert$	$2\lvert E \rvert / \lvert V \rvert$	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	225K	61	3.65	2.99	0.79	0.00027
Power grid	5K	2.67	18.7	12.4	0.08	0.005
C. elegans	0.3L	14	2.65	2.25	0.28	0.05

Small World Models



Watts-Strogatz on 1,000 vertices with rewiring probability $p = 0.05$.

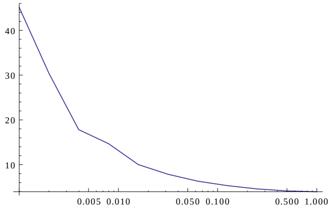
Small World Models

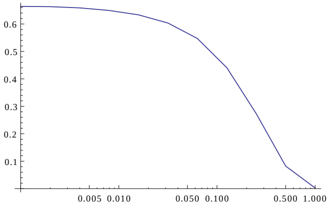


Probability of rewiring, p

Notice the existence of a region of high clustering and low average path length. (e.g. $0.01 \leq p \leq 0.1$)

Small World Models





Average distance Clustering coefficient

Watts-Strogatz graph on 4,000 vertices, starting from a 10-regular graph

- Intuition: if you add a little bit of randomness to a structured graph, you get the small world effect. Related work: See also [Bollobás and Chung, 1988].

Navigation in a Small World

Jon Kleinberg [Kleinberg, 2000]

There exist short paths, but how to find them using only local information?

We will use a simple directed model due to Kleinberg [Kleinberg, 2000].

A local algorithm is an algorithm that based on the following information decides the next vertex on the path:

- can remember the source, the destination and its current location
- and can query the graph to find the long-distance edge at

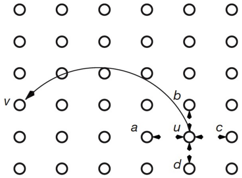
Navigation in a Small World

Define $d(u, v)$ =shortest path distance using only original grid edges.

Directed graph model, parameter r:

- Each vertex is connected to its four adjacent vertices
- For each vertex v we add an extra link (v, u) where u is chosen with probability proportional to $d(v, u)^{-r}$

Notice that compared to the Watts-Strogatz model the long range edges are added in a biased way.



Model (Source [Kleinberg, 2000])

Navigation in a Small World

- When $r = 0$, the end of the long distance edge is uniformly distributed over all vertices independently of the distance.
- As r increases the length of the long distance edges decreases in expectation.

Results

1. $r < 2$: The end points of the long distance edges tend to be uniformly distributed over the vertices of the grid and it is unlikely on a short path to encounter a long distance edge whose end point is close to the destination. No local algorithm can find them.
2. $r = 2$: There are short paths and the simple algorithm that always selects the edge that takes you closest to the destination will find a short path.
3. $r > 2$: With high probability there are no short paths at all.

Navigation in a Small World

- When $r = 0$, the end of the long distance edge is uniformly distributed over all vertices independently of the distance.
- As r increases the length of the long distance edges decreases in expectation.

Results

1. $r < 2$: The end points of the long distance edges tend to be uniformly distributed over the vertices of the grid and it is unlikely on a short path to encounter a long distance edge whose end point is close to the destination. No local algorithm can find them.
2. $r = 2$: There are short paths and the simple algorithm that always selects the edge that takes you closest to the destination will find a short path.
3. $r > 2$: With high probability there are no short paths at all.

Copying model

R. Kumar, P. Raghavan, S. Rajagopalan, D Sivakumar, A. Tomkins and E. Upfal [Kumar et al., 2000] analyze the copying model [Kleinberg et al., 1999b].

- $\alpha \in (0, 1)$: copy factor
- d constant out degree.

Evolving Copying Model, time $t + 1$

• Create a new vertex $t + 1$.
• Choose a prototype vertex $u \in V_t$ uniformly at random.
• The i-th out-link of $t + 1$ is chosen as follows: with probability α we select $x \in V_{t-1}$ uniformly at random and with the remaining probability it copies the i-th out-lin of u.

Copying model

[Kumar et al., 2000] among other theorems, prove that the in-degrees follow a power-law distribution.

Theorem

For $r > 0$ *the limit* $P_r = \lim_{t \to +\infty} \frac{N_t(r)}{t}$ *exists and satisfies*

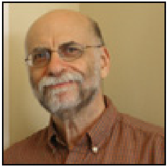
$$P_r = \Theta(r^{-\frac{2-\alpha}{1-\alpha}}).$$

Explains the large number of bipartite cliques in the Web graph.

Static models with power-law degree distributions do not account for this phenomenon!

Cooper-Frieze model

Cooper Frieze



Alan Frieze

Cooper and Frieze [Cooper and Frieze, 2003] introduce a general model

1. with many parameters,
2. that generalizes preferential attachment, generalized preferential attachment and copying models
3. and whose attachment rule is a mixture of preferential and uniform.

Cooper-Frieze model

Some of their findings

1. We can obtain densification and shrinking diameters (add edges among existing vertices)
2. Power law in expectation and strong concentration under mild assumptions.
3. Novel techniques for concentration (martingales+Laplace).

Kronecker graphs

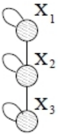
Reminder, Kronecker product

Suppose $A = [a_{ij}]$ is an $m \times n$ matrix and $B = [b_{ij}]$ is a $p \times q$ matrix. Then, $A \otimes B$ is the $mp \times nq$ matrix

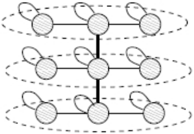
$$\begin{pmatrix} a_{11}B & .. & a_{1n}B \\ .. & .. & .. \\ a_{m1}B & .. & a_{mn}B \end{pmatrix}$$

Leskovec, Chakrabarti, Kleinberg, Faloutsos and Ghahramani [Leskovec et al., 2010] propose a model based on the Kronecker product, generalizing a RMAT [Chakrabarti et al., 2004].

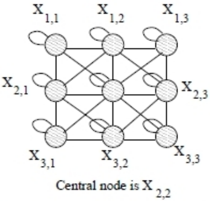
Kronecker graphs



(a) Graph K_1



(b) Intermediate stage



(c) Graph $K_2 = K_1 \otimes K_1$

1	1	0
1	1	1
0	1	1

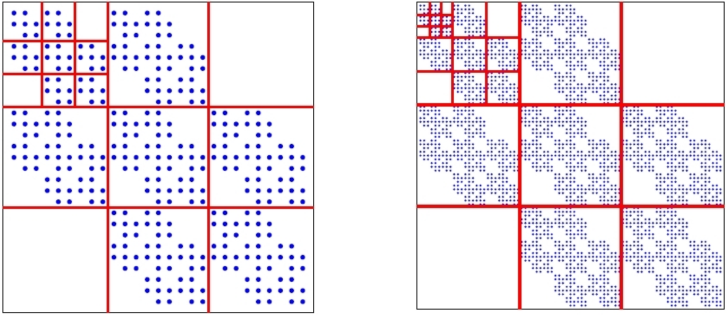
(d) Adjacency matrix of K_1

K_1	K_1	0
K_1	K_1	K_1
0	K_1	K_1

(e) Adjacency matrix of $K_2 = K_1 \otimes K_1$

Source [Leskovec et al., 2010]

Kronecker graphs



(a) K_3 adjacency matrix (27×27) (b) K_4 adjacency matrix (81×81)

Source [Leskovec et al., 2010]

Kronecker graphs

A Stochastic Kronecker graph is defined by two parameters:

- an integer k
- the seed/initiator matrix θ

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

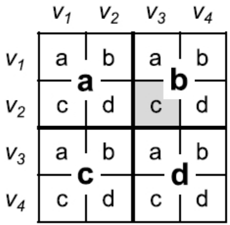
We obtain a graph with $n = 2^k$ vertices by taking repeatedly Kronecker products. Let $A_{k,\theta} = \underbrace{\theta \otimes \ldots \otimes \theta}_{l \text{ times}}$ be the resulting matrix. The graph's adjacency matrix $\bar{A}_{k,\theta}$ is obtained by a randomized rounding A.

- In the vast majority of work so far, 2×2 seed matrices are used. However, noone restricts us from using other seed matrices as well.

Kronecker graphs

	u_1	u_2
u_1	a	b
u_2	c	d

	v_1	v_2	v_3	v_4
v_1	a·a	a·b	b·a	b·b
v_2	a·c	a·d	b·c	b·d
v_3	c·a	c·b	d·a	d·b
v_4	c·c	c·d	d·c	d·d



In practice we never need to compute A, but we can actually do a sampling based on the hierarchical properties of Kronecker products.

Kronecker graphs

Suppose we are given $G(V, E)$ such that $|V| = n = 2^k$.

- Erdös-Rényi

$$\begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

- Core-periphery

$$\begin{pmatrix} 0.9 & 0.5 \\ 0.5 & 0.1 \end{pmatrix}$$

- Hierarchical community structure

$$\begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}$$

Kronecker graphs

- Power law degree distributions [Leskovec et al., 2010]
- Power law eigenvalue distribution [Leskovec et al., 2010]
- Small diameter [Leskovec et al., 2010]
- Densification power law [Leskovec et al., 2010]
- Shrinking diameter [Leskovec et al., 2010]
- Triangles [Tsourakakis, 2008]
- Connectivity [Mahdian and Xu, 2007]
- Giant components [Mahdian and Xu, 2007]
- Diameter [Mahdian and Xu, 2007]
- Searchability [Mahdian and Xu, 2007]

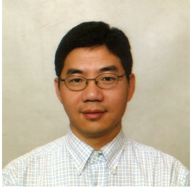
Kronecker graphs

How do we find a seed matrix θ such that $A_G \approx \underbrace{\theta \otimes \ldots \otimes \theta}_{k \text{ times}}$?

- Maximum likelihood estimation: $argmax_\theta \mathbf{Pr}\left[G|\theta\right]$. Hard since exact computation requires $O(n!n^2)$ time, but Metropolis sampling and approximations allow $O(m)$ time good approximations [Leskovec and Faloutsos, 2007].
- Moment based estimation: Express the expected number of certain subgraphs (e.g., edges, triangles, triples) as a function of a, b, c and solve a system of equations [Gleich and Owen, 2012].

Chung-Lu model

Fan Chung Graham



Linyuan Lu

The model is specified by $w = (w_1, \ldots, w_n)$ which represents the expected degree sequence. Vertices i, j are connected with probability

$$p_{ij} = \frac{w_i w_j}{\sum_{k=1}^{n} w_k} = \rho w_i w_j.$$

Clearly, to have a proper probability distribution $w_{max}^2 \leq \rho$. We can obtain an Erdös-Rényi random graph by setting

$$w = (pn, \ldots, pn).$$

Chung-Lu model

How to set the weights to get power law exponent β? In the power law model the probability of having degree k:

$$\mathbf{Pr}\left[deg(v) = k\right] = \frac{k^{-\beta}}{\zeta(\beta)}.$$

Hence, for $\beta > 1$

$$\mathbf{Pr}\left[deg(v) \geq k\right] = \sum_{l \geq k}^{+\infty} \frac{k^{-\beta}}{\zeta(\beta)} = \frac{1}{\zeta(\beta)(\beta - 1)k^{\beta - 1}}.$$

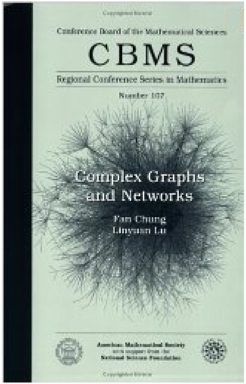
Assuming the weights are decreasing and setting $w_i = k, i/n = \mathbf{Pr}\left[deg(v) \geq k\right]$.

$$w_i = \left(\frac{i}{\zeta(\beta)(\beta - 1)i}\right)^{-\frac{1}{\beta - 1}}.$$

Chung-Lu model

Rigorous results on:

- Degree sequence
- Giant component
- Average distance and the diameter
- Eigenvalues of the adjacency and the Laplacian matrix
- ...



Kronecker vs. Chung-Lu

Pinar, Seshahdri, Kolda [Pinar et al., 2011] compare the two models and find "the SKG model is close enough to its associated CL model that most users of SKG could just as well use the CL model for generating graphs."

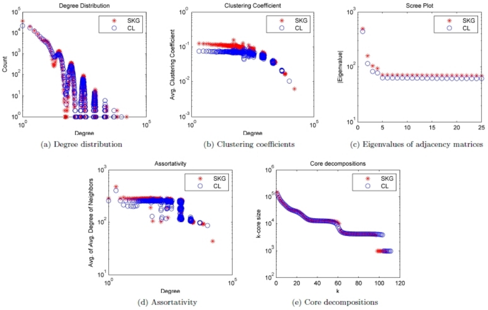


Figure 1: Comparison of the graph properties of SKG generated with Graph500 parameters and an equivalent CL.

Source [Pinar et al., 2011]

Forest-fire model

J. Leskovec

J. Kleinberg

C. Faloutsos

Leskovec, Kleinberg and Faloutsos [Leskovec et al., 2007] propose the forest fire model that is able to re-produce at a qualitative scale most of the established properties of real-world networks.

Forest-fire model

Basic version of the model

1. p : forward burning probability
2. r : backward burning ratio

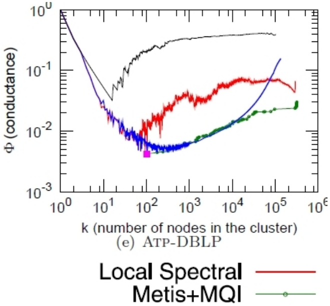
- Initially, we have a single vertex
- At time t a new vertex v arrives to G_t
  - Node v picks an *ambassador/seed* node u uniformly at random link to u.
  - Two numbers x, y are sampled from two geometric distributions with parameters $\frac{p}{1-p}$ and $\frac{rp}{1-rp}$ respectively. Then, v chooses x out-links and y in-links of u which are incident to unvisited vertices. Let $u_1, \ldots, u_{x+y}$ be these chosen endpoints.
  - Mark $u_1, \ldots, u_{x+y}$ as visited and apply the previous step recursively to each of them.

Forest-fire model

The forest-fire model is able to

- Heavy tailed in-degrees and out-degrees
- Densification power law
- Shrinking diameter
- ...
- Deep cuts at small size scales and the absence of deep cuts at large size scales

Reminder



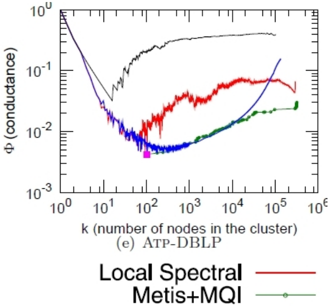
(e) ATP-DBLP

Forest-fire model

The forest-fire model is able to

- Heavy tailed in-degrees and out-degrees
- Densification power law
- Shrinking diameter
- ...
- Deep cuts at small size scales and the absence of deep cuts at large size scales

Reminder



(e) ATP-DBLP

Random Graphs Applications

Influence of Search Engines on Preferential Attachment

Junghoo Cho, UCLA

Search-Engine Bias Project

- In their early days, search engines merely observed and exploited the Web graph for ranking.
- Nowadays, they are unquestionably influencing the evolution of the Web graph.
- How?

Influence of Search Engines on Preferential Attachment

- "Virtuous circle of limelight": A search engine ranks a page highly → Web page owners find this page more often and link to it → raises its popularity and so on...
- Main finding: Cho and his collaborators [Cho and Roy, 2004] estimate that the time taken for a page to reach prominence can be delayed by a factor of over 60 if a search engine diverts clicks to popular pages.
- We use random graphs to obtain insights into this phenomenon [Chakrabarti et al., 2005].

Influence of Search Engines on Preferential Attachment

- "Virtuous circle of limelight": A search engine ranks a page highly → Web page owners find this page more often and link to it → raises its popularity and so on...
- Main finding: Cho and his collaborators [Cho and Roy, 2004] estimate that the time taken for a page to reach prominence can be delayed by a factor of over 60 if a search engine diverts clicks to popular pages.
- We use random graphs to obtain insights into this phenomenon [Chakrabarti et al., 2005].

Influence of Search Engines on Preferential Attachment

- "Virtuous circle of limelight": A search engine ranks a page highly → Web page owners find this page more often and link to it → raises its popularity and so on...
- Main finding: Cho and his collaborators [Cho and Roy, 2004] estimate that the time taken for a page to reach prominence can be delayed by a factor of over 60 if a search engine diverts clicks to popular pages.
- We use random graphs to obtain insights into this phenomenon [Chakrabarti et al., 2005].

Influence of Search Engines on Preferential Attachment

Frieze, Vera and Chakrabarti [Chakrabarti et al., 2005] introduce a model with two parameters:

- p, a probability
- N, maximum number of celebrity nodes listed by the search engine
- m, edge parameter

Notation:

- Sequence of graphs $\{G_t\}_{t=1}^{+\infty}$. G_t will have t vertices and mt edges.
- $D_t(U) = \sum_{x \in U} deg_t(x)$
- S_t the set of at most N vertices with largest degrees in G_t.
- $d_k(t)$ denotes the number of vertices of degree k at time t. in the set $V_t - S_t$.

Influence of Search Engines on Preferential Attachment

- **Time step 1:** The process is initialized with graph G_1 which consists of an isolated vertex x_1 and m loops.
- **Time step $t > 1$:** We add a vertex x_t to G_{t-1}. We then add m random edges (x_t, y_i), $i = 1, \ldots, m$ incident with x_t, where y_i are nodes in G_{t-1}. For each i:
  - With probability p we choose $y_i \in S_{t-1}$.
  - With probability $1 - p$ we choose $y_i \in V_{t-1}$.

  In both case y_i is selected by preferential attachment, i.e.,

  $$\mathbf{Pr}\left[y_i = x\right] = \frac{deg_{t-1}(x)}{\sum_{u \in U} deg_{t-1}(u)},$$

  where $U = S_{t-1}$ or $U = V_{t-1}$.

Influence of Search Engines on Preferential Attachment

Theorem

Let $m \geq \max\{15, \frac{2}{1-p}\}$ *and* $0 < p < 1$

- *Let* $S_t = \{s_1, \ldots s_N\}$ *in decreasing order of degree. Then* $\mathbb{E}\left[deg_t(s_i)\right] \sim \alpha_i t$ *for every* $i \leq N$ *for some constant* $\alpha_i > 0$.
- *There is an absolute constant* A_1 *such that for every* $k \geq m$

$$\mathbb{E}\left[d_k(t)\right] = \frac{A_1 n}{k^{1+\frac{2}{1-p}}} + \textit{second order terms.}$$

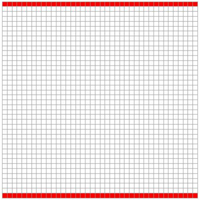
Influence of Search Engines on Preferential Attachment

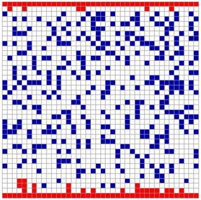
The theorem and its proof verify our intuition.

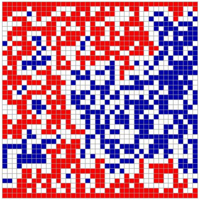
- The celebrity lit gets fixed quickly.
- Each celebrity page captures a constant fraction of all edges ever generated in the graph.
- The non-celebrity vertices obey a power law which is steeper.

Robustness and Vulnerability

- Intuitively, a complex network is *robust* if it keeps its basic functionality under the failure of some of its components.
- We distinguish between random failure and intentional attacks.
- Related to percolation



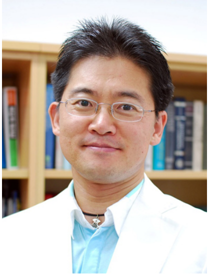




Percolation

Robustness and Vulnerability

R. Albert



H. Jeong

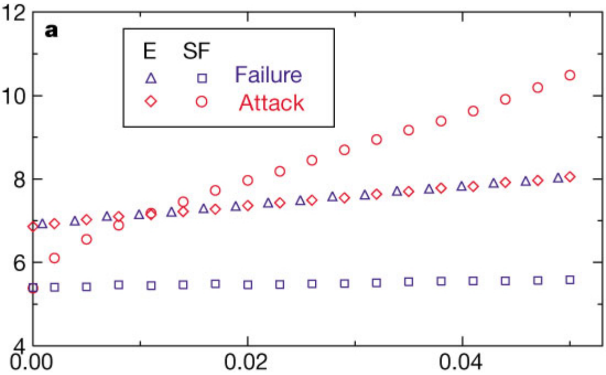
L. Barabási

Albert, Jeong and Barabási [Albert et al., 2000] provide simulations which indicate that scale free networks are robust to random failures.

10 second sound bite science

The Internet is robust yet fragile. 95% of the links can be removed and the graph will stay connected. However, targeted removal of 2.3% of the hubs would disconnect the Internet.

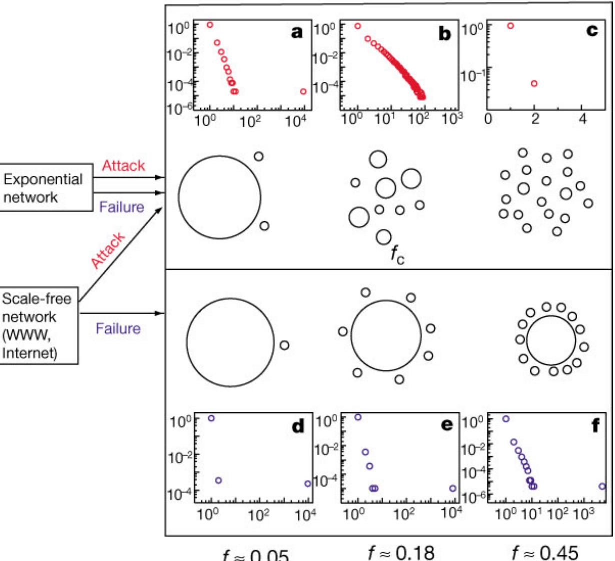
Robustness and Vulnerability



$n = 10\,000$ vertices $m = 20\,000$ links

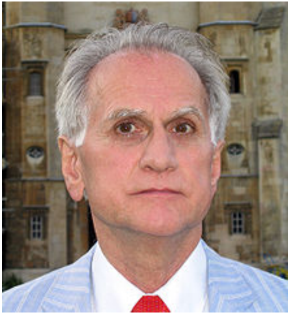
- Diameter of an Erdös-Rényi and a scale-free network as a function of the fraction f of vertices deleted.
- The power-law distribution implies that under random sampling, vertices with small degree are selected selected with much higher probability.

Robustness and Vulnerability



- The cluster size distribution for various values of f for an Erdös-Rényi graph and a scale-free network under random and malicious failures (Source [Albert et al., 2000]).
- Intuition: Scale-free graphs are inhomogeneous and this implies both better performance under random failures and reduced attack survivability.

Robustness and Vulnerability



Béla Bollobás

Oliver Riordan

Bollobás and Riordan [Bollobás and Riordan, 2004] studied the robustness and vulnerability of a scale-free graph, using specifically the Barabási-Albert model.

Robustness and Vulnerability

- When vertices of $G_m^{(n)}$ are deleted independently with probability $1-p$, there is always a giant component! In other words, there is no critical p.
- However the size of the 'giant' component depends on p.

Theorem

Let $m \geq 2$, $0 < p < 1$ be fixed and let G_p be obtained from $G_m^{(n)}$ by deleting vertices independently with probability $1-p$. Then as $n \to +\infty$ **whp** *the largest component of G_p has order $((c(p,m)+o(1))n$. Furthermore, as $p \to 0$ with m fixed, $c(p,m) = \exp\left(\frac{1}{O(p)}\right)$.*

Robustness and Vulnerability

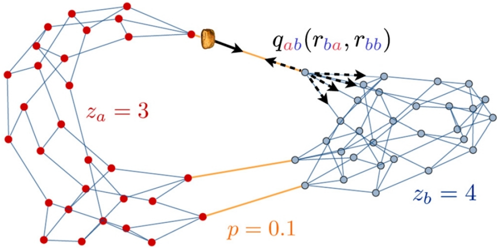
- When $G_m^{(n)}$ is deliberately attacked, finding the "best" attack is hard.
- Bollobás and Riordan consider the natural attack of deleting the earliest vertices up to some cutoff cn.

Theorem

Let G_c be obtained by $G_m^{(n)}$ by deleting all vertices with index less than cn, where $0 < c < 1$ is a constant. Let $c_m = \frac{m-1}{m+1}$. If $c < c_m$ then **whp** *G_c has a component with $\Theta(n)$ vertices. If $c > c_m$ then* **whp** *G_c has no such component.*

More applications

- Study of interdependent networks [Brummitt et al., 2012].



A random three- and four-regular graph connected by Bernoulli distributed coupling with interconnectivity parameter $p = 0.1$ (Source [Brummitt et al., 2012]).

More applications

Itai Ashlagi

Alvin Roth

Compatibility graph : each vertex is a donor-patient pair and each edge between two vertices denotes compatibility for kidney exchange.

- Model kidney exchange with many patient-donor pairs as a random compatibility graph

More applications

Motivation

- We wish to messages in a cellular network G, between any two vertices in a pipeline.
- We require that each link on the route between the vertices (namely, each edge on the path) is assigned a distinct channel (e.g., a distinct frequency).

An edge colored graph G is rainbow edge connected if any two vertices are connected by a path whose edges have distinct colors.

Goal: Find the minimum number of colors needed to rainbow color the edges of G.

Frieze and CET [Frieze and Tsourakakis, 2012] study *rainbow connectivity* in sparse random graphs.

More applications

- Cooper and Frieze [Cooper and Frieze, 2004] studied the performance of crawlers in random evolving scale-free graphs.
- Valiant [Valiant, 2005] uses random graphs to model memorization and association functionalities of the brain.
- Simulations (epidemics, performance of algorithms etc.)
- Graph anonymization [Leskovec et al., 2005a].
- Allow us to argue about the structure of real-world networks. For instance, given a random graph with a fixed degree distribution, what do we expect for the spectrum, subgraphs etc?
- Give rise to objectives by using them as null models (modularity)
- and many more ..

Conclusions

- Graphs are simple models that can represent a wide variety of datasets
- Real graphs from different application domains have many similar properties but also differences

Conclusions

- Just scratched the tip of the iceberg

acknowledgements

Slides come from the **Algorithmic techniques for modeling and mining large graphs** tutorial

with Alan Frieze, Aristides Gionis that was presented at KDD 2013 and ECML PKDD 2013

references I

(2000).
The web is a bow tie.
Nature, 405.

Albert, R., Jeong, H., and Barabási, A.-L. (2000).
Error and attack tolerance of complex networks.
Nature, 406(6794):378–382.

Bollobás, B. and Chung, F. R. K. (1988).
The diameter of a cycle plus a random matching.
SIAM Journal on discrete mathematics, 1(3):328–333.

Bollobás, B. and Riordan, O. (2003).
Mathematical results on scale-free random graphs.
Handbook of graphs and networks, 1:34.

references II

Bollobás, B. and Riordan, O. (2004).
Robustness and vulnerability of scale-free random graphs.
Internet Mathematics, 1(1):1–35.

Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., and Wiener, J. (2000).
Graph structure in the web: Experiments and models.
In *Proceedings of the Ninth Conference on World Wide Web*, pages 309–320, Amsterdam, Netherlands. ACM Press.

Brummitt, C. D., DSouza, R. M., and Leicht, E. (2012).
Suppressing cascades of load in interdependent networks.
Proceedings of the National Academy of Sciences, 109(12):E680–E689.

references III

Buckley, P. G. and Osthus, D. (2004).

Popularity based random graph models leading to a scale-free degree sequence.

Discrete Mathematics, 282(1):53–68.

Callaway, D. S., Hopcroft, J. E., Kleinberg, J. M., Newman, M. E., and Strogatz, S. H. (2001).

Are randomly grown graphs really random?

Physical Review E, 64(4):041902.

Chakrabarti, D., Zhan, Y., and Faloutsos, C. (2004).

R-mat: A recursive model for graph mining.

Computer Science Department, page 541.

references IV

Chakrabarti, S., Frieze, A., and Vera, J. (2005).
The influence of search engines on preferential attachment.
In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 293–300. Society for Industrial and Applied Mathematics.

Cho, J. and Roy, S. (2004).
Impact of search engines on page popularity.
In *Proceedings of the 13th international conference on World Wide Web*, pages 20–29. ACM.

Clauset, A., Shalizi, C. R., and Newman, M. E. (2009).
Power-law distributions in empirical data.
SIAM review, 51(4):661–703.

references V

Cooper, C. and Frieze, A. (2003).
A general model of web graphs.
Random Structures & Algorithms, 22(3):311–335.

Cooper, C. and Frieze, A. (2004).
Crawling on simple models of web graphs.
Internet Mathematics, 1(1):57–90.

Dorogovtsev, S. N. and Mendes, J. F. (2002).
Evolution of networks.
Advances in physics, 51(4):1079–1187.

Frieze, A. and Tsourakakis, C. E. (2012).
Rainbow connection of sparse random graphs.
the electronic journal of combinatorics, 19(4):P5.

references VI

Gleich, D. F. and Owen, A. B. (2012).
Moment-based estimation of stochastic kronecker graph parameters.
Internet Mathematics, 8(3):232–256.

Holme, P. and Kim, B. J. (2002).
Growing scale-free networks with tunable clustering.
Physical Review E, 65(2):026107.

Kleinberg, J. M. (2000).
Navigation in a small world.
Nature, 406(6798):845–845.

Kleinberg, J. M., Kumar, R., Raghavan, P., Rajagopalan, S., and Tomkins, A. S. (1999a).
The web as a graph: Measurements, models, and methods.
In *Computing and combinatorics*, pages 1–17. Springer.

references VII

Kleinberg, J. M., Kumar, R., Raghavan, P., Rajagopalan, S., and Tomkins, A. S. (1999b).

The Web as a graph: measurements, models and methods.

In *Proceedings of the 5th Annual International Computing and Combinatorics Conference (COCOON)*, volume 1627 of *Lecture Notes in Computer Science*, pages 1–18, Tokyo, Japan. Springer.

Krivelevich, M. and Sudakov, B. (2013).

The phase transition in random graphs - a simple proof.

Kumar, R., Raghavan, P., Rajagopalan, S., Sivakumar, D., Tomkins, A., and Upfal, E. (2000).

Stochastic models for the web graph.

In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 57–65, Redondo Beach, CA, USA. IEEE CS Press.

Kumar, R., Raghavan, P., Rajagopalan, S., and Tomkins, A. (1999).
Trawling the Web for emerging cyber-communities.
Computer Networks, 31(11–16):1481–1493.

Lakhina, A., Byers, J. W., Crovella, M., and Xie, P. (2003).
Sampling biases in ip topology measurements.
In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 1, pages 332–341. IEEE.

Leskovec, J., Chakrabarti, D., Kleinberg, J., and Faloutsos, C. (2005a).
Realistic, mathematically tractable graph generation and evolution, using kronecker multiplication.
In *Knowledge Discovery in Databases: PKDD 2005*, pages 133–145. Springer.

references IX

Leskovec, J., Chakrabarti, D., Kleinberg, J., Faloutsos, C., and Ghahramani, Z. (2010).
Kronecker graphs: An approach to modeling networks.
The Journal of Machine Learning Research, 11:985–1042.

Leskovec, J. and Faloutsos, C. (2007).
Scalable modeling of real graphs using kronecker multiplication.
In *Proceedings of the 24th international conference on Machine learning*, pages 497–504. ACM.

Leskovec, J., Kleinberg, J., and Faloutsos, C. (2005b).
Graphs over time: densification laws, shrinking diameters and possible explanations.
In *KDD '05: Proceeding of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 177–187, New York, NY, USA. ACM Press.

references X

Leskovec, J., Kleinberg, J., and Faloutsos, C. (2007).
Graph evolution: Densification and shrinking diameters.
ACM Transactions on Knowledge Discovery from Data (TKDD), 1(1):2.

Leskovec, J., Lang, K. J., Dasgupta, A., and Mahoney, M. W. (2009).
Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters.
Internet Mathematics, 6(1):29–123.

M. Faloutsos, P. Faloutsos, C. F. (1999).
On power-law relationships of the internet topology.
In *SIGCOMM*.

references XI

Mahdian, M. and Xu, Y. (2007).

Stochastic kronecker graphs.

In *Algorithms and models for the web-graph*, pages 179–186. Springer.

McKay, B. D. and Wormald, N. C. (1997).

The degree sequence of a random graph. i. the models.

Random Structures and Algorithms, 11(2):97–117.

Newman, M. E. J. (2003).

The structure and function of complex networks.

Ostroumova, L., Ryabchenko, A., and Samosvat, E. (2012).

Generalized preferential attachment: tunable power-law degree distribution and clustering coefficient.

arXiv preprint arXiv:1205.3015.

references XII

Pinar, A., Seshadhri, C., and Kolda, T. G. (2011).

The similarity between stochastic kronecker and chung-lu graph models.

arXiv preprint arXiv:1110.4925.

Tsourakakis, C. E. (2008).

Fast counting of triangles in large real networks without counting: Algorithms and laws.

In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, pages 608–617. IEEE.

Valiant, L. G. (2005).

Memorization and association on a realistic neural model.

Neural computation, 17(3):527–555.