The science of networks

(1) Real world networks and their universal properties
(2) Mathematical framework
(3) Networks of configuration type
(4) Preferential attachment networks

Real world networks and their universal properties

Networks consists of a large, but finite, number of nodes connected by links.
In the modern world, networks are ubiquitous:

- social and communication networks,
- world wide web and internet,
- scientific and other collaboration graphs, . . .

Real world networks and their universal properties

Networks consists of a large, but finite, number of nodes connected by links.
In the modern world, networks are ubiquitous:

- social and communication networks,
- world wide web and internet,
- scientific and other collaboration graphs, ...

Scientists who have studied these networks have made a number of surprising discoveries, based on statistical, numerical and nonrigorous analytical arguments.
Our aim is to put these claims in a rigorous mathematical framework and to verify or refine them in this framework.

Real world networks and their universal properties

A crucial claim often made is that these networks are scale-free. In the *language of network science* this means that

$$\frac{\#\text{nodes with degree } k}{\#\text{nodes in the network}} \approx k^{-\tau},$$

when k is large. Note that this is not a mathematically rigorous definition. The parameter τ is called the power-law exponent.

Real world networks and their universal properties

A crucial claim often made is that these networks are scale-free. In the *language of network science* this means that

$$\frac{\#\text{nodes with degree } k}{\#\text{nodes in the network}} \approx k^{-\tau},$$

when k is large. Note that this is not a mathematically rigorous definition. The parameter τ is called the power-law exponent.

A claim often made is that of universality. This means that many global features of the network only depend on a few observable parameters, in particular the power-law exponent.

Real world networks and their universal properties

A crucial claim often made is that these networks are scale-free. In the *language of network science* this means that

$$\frac{\#\text{nodes with degree } k}{\#\text{nodes in the network}} \approx k^{-\tau},$$

when k is large. Note that this is not a mathematically rigorous definition. The parameter τ is called the power-law exponent.

A claim often made is that of universality. This means that many global features of the network only depend on a few observable parameters, in particular the power-law exponent.

Some estimated power-law exponents:

- an e-mail network at the University of Kiel: $\tau \approx 1.81$,
- the world-wide web: $\tau \approx 2.1$,
- the internet: $\tau \approx 2.2$,
- the movie actor network: $\tau \approx 2.3$,
- the collaboration graph in mathematics: $\tau \approx 2.4$,
- a network of sexual relationships in Sweden: $\tau \approx 3.3$.

Real world networks and their universal properties

The networks we are interested in are not necessarily connected in the mathematical sense, but typically there is a large connected subnetwork. In the *language of network science* a network has a giant component if its largest connected component comprises a positive fraction of all nodes.

Real world networks and their universal properties

The networks we are interested in are not necessarily connected in the mathematical sense, but typically there is a large connected subnetwork. In the *language of network science* a network has a giant component if its largest connected component comprises a positive fraction of all nodes.

Real networks may or may not be robust, which in the *language of network science* means that after removal of an arbitrary fraction $0 < p < 1$ of randomly chosen links it still has a giant component.

Real world networks and their universal properties

The networks we are interested in are not necessarily connected in the mathematical sense, but typically there is a large connected subnetwork. In the *language of network science* a network has a giant component if its largest connected component comprises a positive fraction of all nodes.

Real networks may or may not be robust, which in the *language of network science* means that after removal of an arbitrary fraction $0 < p < 1$ of randomly chosen links it still has a giant component.

Claim 1

Networks are robust iff $\tau \leq 3$.

Real world networks and their universal properties

The networks we are interested in are not necessarily connected in the mathematical sense, but typically there is a large connected subnetwork. In the *language of network science* a network has a giant component if its largest connected component comprises a positive fraction of all nodes.

Real networks may or may not be robust, which in the *language of network science* means that after removal of an arbitrary fraction $0 < p < 1$ of randomly chosen links it still has a giant component.

Claim 1

Networks are robust iff $\tau \leq 3$.

Real networks may or may not be vulnerable, which in the *language of network science* means that after targeted removal of a small number of carefully chosen links the giant component can be destroyed.

Real world networks and their universal properties

The networks we are interested in are not necessarily connected in the mathematical sense, but typically there is a large connected subnetwork. In the *language of network science* a network has a giant component if its largest connected component comprises a positive fraction of all nodes.

Real networks may or may not be robust, which in the *language of network science* means that after removal of an arbitrary fraction $0 < p < 1$ of randomly chosen links it still has a giant component.

Claim 1

Networks are robust iff $\tau \leq 3$.

Real networks may or may not be vulnerable, which in the *language of network science* means that after targeted removal of a small number of carefully chosen links the giant component can be destroyed.

Claim 2

Networks are vulnerable iff $\tau \leq 3$.

Real world networks and their universal properties

The typical distances observed in networks are often very small. In the *language of network science* networks are called small worlds if the distance of two randomly chosen nodes in the giant component is of order $\log N$, and they are called ultrasmall if it is of order $\log \log N$, where N is the total number of nodes.

Real world networks and their universal properties

The typical distances observed in networks are often very small. In the *language of network science* networks are called small worlds if the distance of two randomly chosen nodes in the giant component is of order $\log N$, and they are called ultrasmall if it is of order $\log \log N$, where N is the total number of nodes.

Claim 3

Networks are ultrasmall iff $\tau \in (2, 3)$.

Real world networks and their universal properties

The typical distances observed in networks are often very small. In the *language of network science* networks are called small worlds if the distance of two randomly chosen nodes in the giant component is of order $\log N$, and they are called ultrasmall if it is of order $\log \log N$, where N is the total number of nodes.

Claim 3

Networks are ultrasmall iff $\tau \in (2, 3)$.

There are a lot of further claims in the network sciences literature, typically about processes on the networks but we stop here and start with our journey looking at mathematical models and results.

Real world networks and their universal properties

A small selection of references:

- Albert, Jeong, Barabasi. Error and attack tolerance of complex networks. Nature 406, 378-382 (2000)
- Albert, Barabasi. Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 4797 (2002)
- Cohen, Havlin. Scale-free networks are ultrasmall. Phys. Rev. Lett. 90, 058701 (2003)
- Dorogovtsev, Mendes. Evolution of Networks. Adv. Phys. 51, 1079-1187 (2002)
- Newman. The structure and function of complex networks. SIAM Review 45, 167-256 (2003)

Mathematical framework

Because of the large size of real networks random models are particularly suitable to study their behaviour and to test and refine the claims made.

Mathematical framework

Because of the large size of real networks random models are particularly suitable to study their behaviour and to test and refine the claims made.

The general approach is to define a sequence $(\mathcal{G}_N)$ of random graphs with N vertices and study asymptotic properties as N goes to infinity. In this framework we can give rigorous definitions of the main notions of network science.

Mathematical framework

Because of the large size of real networks random models are particularly suitable to study their behaviour and to test and refine the claims made.

The general approach is to define a sequence $(\mathcal{G}_N)$ of random graphs with N vertices and study asymptotic properties as N goes to infinity. In this framework we can give rigorous definitions of the main notions of network science.

We always assume that the vertices of $\mathcal{G}_N$ are labelled as $1, \ldots, N$ and define the empirical degree distribution of $(\mathcal{G}_N)$ as $(X_N(k) \colon k = 0, 1, \ldots)$ where

$$X_N(k) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{\text{degree of vertex } i = k\}.$$

We call $(\mathcal{G}_N)$ scale-free with power-law exponent τ if

$$\lim_{N \to \infty} X_N(k) = \mu(k) \quad \text{in probability,}$$

for some nonrandom probability vector $(\mu(k) \colon k = 0, 1, \ldots)$ and

$$\lim_{k \to \infty} \frac{\log \mu(k)}{\log k} = -\tau.$$

Mathematical framework

Let $\mathcal{C}_N \subset \mathcal{G}_N$ be the largest connected component of the network. We say that $(\mathcal{G}_N)$ has a giant component if

$$\lim_{N\to\infty} \frac{\#\mathcal{C}_N}{N} = p > 0 \quad \text{in probability.}$$

Mathematical framework

Let $\mathcal{C}_N \subset \mathcal{G}_N$ be the largest connected component of the network. We say that $(\mathcal{G}_N)$ has a giant component if

$$\lim_{N\to\infty} \frac{\#\mathcal{C}_N}{N} = p > 0 \quad \text{in probability.}$$

Given $\mathcal{G}_N$ and a deletion parameter q we obtain the percolated network $\mathcal{G}_N(q)$ by removing every edge of $\mathcal{G}_N$ independently with probability q. We say the network is robust if, for every $0 < q < 1$ the network $(\mathcal{G}_N(q))$ has a giant component.

Mathematical framework

Let $\mathcal{C}_N \subset \mathcal{G}_N$ be the largest connected component of the network. We say that $(\mathcal{G}_N)$ has a giant component if

$$\lim_{N\to\infty} \frac{\#\mathcal{C}_N}{N} = p > 0 \quad \text{in probability.}$$

Given $\mathcal{G}_N$ and a deletion parameter q we obtain the percolated network $\mathcal{G}_N(q)$ by removing every edge of $\mathcal{G}_N$ independently with probability q. We say the network is robust if, for every $0 < q < 1$ the network $(\mathcal{G}_N(q))$ has a giant component.

Given $\mathcal{G}_N$ we let $d(\cdot\,,\,\cdot)$ be the graph distance of two vertices, i.e. the length of the shortest path between them. Picking two vertices $V, W \in \mathcal{C}_N$ independently, uniformly from $\mathcal{C}_N$, we say the network is ultrasmall if

$$\lim_{N\to\infty} \frac{d(V,W)}{\log\log N} = c > 0 \quad \text{in probability.}$$

Networks of configuration type

The most studied network model in the mathematical literature is the sparse Erdős-Rényi graph . Given a parameter $p > 0$ we obtain $\mathcal{G}_N$ by putting an edge between any two vertices independently with probability p/N.

Networks of configuration type

The most studied network model in the mathematical literature is the sparse Erdős-Rényi graph . Given a parameter $p > 0$ we obtain $\mathcal{G}_N$ by putting an edge between any two vertices independently with probability p/N.

In this model the degree of a vertex is binomially distributed with parameters $N-1$ and p/N, so that by the law of small numbers

$$\lim_{N\to\infty} X_N(k) = e^{-p}\frac{p^k}{k!} \quad \text{in probability.}$$

Networks of configuration type

The most studied network model in the mathematical literature is the sparse Erdős-Rényi graph . Given a parameter $p > 0$ we obtain $\mathcal{G}_N$ by putting an edge between any two vertices independently with probability p/N.

In this model the degree of a vertex is binomially distributed with parameters $N - 1$ and p/N, so that by the law of small numbers

$$\lim_{N\to\infty} X_N(k) = e^{-p}\frac{p^k}{k!} \qquad \text{in probability.}$$

In particular these networks are not scale free as the asymptotic degree distribution has light tails. It is not a suitable model for 'real' networks.

Networks of configuration type

A first alternative is to consider networks with fixed degree sequence, as studied for example by Bollobas (1980) and Aiello, Chung, Lu (2001).

Networks of configuration type

A first alternative is to consider networks with fixed degree sequence, as studied for example by Bollobas (1980) and Aiello, Chung, Lu (2001).

Take $D_1, D_2, \ldots$ iid random variables with

$$\mathbb{P}\{D_1 > x\} = x^{1-\tau}(c + o(1)) \qquad \text{as } x \uparrow \infty$$

and values in the nonnegative integers. This sequence will (almost) be the degree sequence in our network.

Networks of configuration type

A first alternative is to consider networks with fixed degree sequence, as studied for example by Bollobas (1980) and Aiello, Chung, Lu (2001).

Take $D_1, D_2, \dots$ iid random variables with

$$\mathbb{P}\{D_1 > x\} = x^{1-\tau}(c + o(1)) \quad \text{as } x \uparrow \infty$$

and values in the nonnegative integers. This sequence will (almost) be the degree sequence in our network.

Given $D_1, \dots, D_N$ we construct the network $\mathcal{G}_N$ as follows:

- To any vertex $n \leq N$ we attach D_n half-edges or stubs.
- We start by matching a stub with a (uniformly) randomly chosen other stub, and continue matching every unpaired stub with a remaining randomly chosen stub until all (or all but one) stubs are matched.
- Any matched pair of stubs are connected to form an edge.

Networks of configuration type

A first alternative is to consider networks with fixed degree sequence, as studied for example by Bollobas (1980) and Aiello, Chung, Lu (2001).

Take $D_1, D_2, \ldots$ iid random variables with

$$\mathbb{P}\{D_1 > x\} = x^{1-\tau}(c + o(1)) \quad \text{as } x \uparrow \infty$$

and values in the nonnegative integers. This sequence will (almost) be the degree sequence in our network.

Given $D_1, \ldots, D_N$ we construct the network $\mathcal{G}_N$ as follows:

- To any vertex $n \leq N$ we attach D_n half-edges or stubs.
- We start by matching a stub with a (uniformly) randomly chosen other stub, and continue matching every unpaired stub with a remaining randomly chosen stub until all (or all but one) stubs are matched.
- Any matched pair of stubs are connected to form an edge.

Obviously the resulting network can have self-loops and multiple edges. However, if $\tau > 2$, the network has power law exponent τ even if self-loops and multiple edges are removed.

Networks of configuration type

An alternative to fixing the degree sequence, is to introduce a fitness for any vertex, and creating an edge between vertices with a probability proportional to the product of their fitnesses.

Networks of configuration type

An alternative to fixing the degree sequence, is to introduce a fitness for any vertex, and creating an edge between vertices with a probability proportional to the product of their fitnesses.

We describe a model introduced by Norros and Reittu under the name conditionally Poissonian random graph. It is based on drawing an iid fitness sequence $\Lambda_1, \Lambda_2, \ldots$ with

$$\mathbb{P}\{\Lambda_1 > x\} = x^{1-\tau}(c + o(1)) \qquad \text{as } x \uparrow \infty$$

Conditional on this sequence, the network is constructed as follows:

Networks of configuration type

An alternative to fixing the degree sequence, is to introduce a fitness for any vertex, and creating an edge between vertices with a probability proportional to the product of their fitnesses.

We describe a model introduced by Norros and Reittu under the name conditionally Poissonian random graph. It is based on drawing an iid fitness sequence $\Lambda_1, \Lambda_2, \ldots$ with

$$\mathbb{P}\{\Lambda_1 > x\} = x^{1-\tau}(c + o(1)) \quad \text{as } x \uparrow \infty$$

Conditional on this sequence, the network is constructed as follows:

- $\mathcal{G}_1$ consists of a single vertex and no edges,
- given $\mathcal{G}_N$ we insert one new vertex and, independently for any $n \leq N$ introduce a random number of edges between the new vertex and n according to a Poisson distribution with mean
$$\frac{\Lambda_n \Lambda_{N+1}}{\sum_{k=1}^{N+1} \Lambda_k},$$
- we further remove each edge in $\mathcal{G}_N$ independently with probability
$$1 - \frac{\sum_{k=1}^{N} \Lambda_k}{\sum_{k=1}^{N+1} \Lambda_k},$$

and thus obtain $\mathcal{G}_{N+1}$.

Networks of configuration type

The conditionally Poissonian random graph is scale-free with power-law exponent τ. If $\tau > 2$ this remains true when multiple edges and self-loops are removed.

The conditionally Poissonian random graph is scale-free with power-law exponent τ. If $\tau > 2$ this remains true when multiple edges and self-loops are removed.

The claims of network science can be investigated for the models of configuration type and to some extent this has been done. But the main criticism is that these models do not explain why real networks are scale-free.

Networks of configuration type

The conditionally Poissonian random graph is scale-free with power-law exponent τ. If $\tau > 2$ this remains true when multiple edges and self-loops are removed.

The claims of network science can be investigated for the models of configuration type and to some extent this has been done. But the main criticism is that these models do not explain why real networks are scale-free.

The preferential attachment paradigm claims to offer a simple and credible explanation for the occurrence of scale-free networks. At the same time it gives rise to a very nice class of network models, which can still be studied rigorously, although they are more complex than the configuration type models.

Networks of configuration type

A small selection of references:

- Aiello, Chung, Lu. A random graph model for power law graphs. Experiment. Math. 10, 53-66 (2001)
- Britton, Deijfen, Martin-Löf . Generating simple random graphs with prescribed degree distribution. J. Statist. Phys. 124, 1377–1397 (2006).
- van der Hofstad, Hooghiemstra. Universality for distances in power-law random graphs. J. Math. Phys. 49, 125-209 (2008).
- Norros, Reittu. On a conditionally Poissonian graph process. Adv. in Appl. Probab. 38, 59-75 (2006)
- Norros, Reittu. Attack resistance of power-law random graphs. Internet Math. 5, 251-266 (2008)

Preferential attachment networks

Popularised by Barabási and Albert (1999) preferential attachment networks are dynamic network models using the current degree of a vertex as its fitness.

Preferential attachment networks

Popularised by Barabási and Albert (1999) preferential attachment networks are dynamic network models using the current degree of a vertex as its fitness.

Growing networks are built by adding vertices successively. When a new vertex is introduced, attachment to vertices with higher degree is preferred, following the principle that the rich get richer. Roughly speaking, a new vertex is connected by edges to a fixed or random number of existing nodes with a probability proportional to a nondecreasing function f of their degree. The function f, which regulates the strength of the preferential attachment is called the attachment rule.

Preferential attachment networks

Popularised by Barabási and Albert (1999) preferential attachment networks are dynamic network models using the current degree of a vertex as its fitness.

Growing networks are built by adding vertices successively. When a new vertex is introduced, attachment to vertices with higher degree is preferred, following the principle that the rich get richer. Roughly speaking, a new vertex is connected by edges to a fixed or random number of existing nodes with a probability proportional to a nondecreasing function f of their degree. The function f, which regulates the strength of the preferential attachment is called the attachment rule.

We first dicuss a version of the model where new vertices are connected to a fixed number $m \geq 2$ of old vertices. Here the attachment rule is affine, more precisely there exist $\delta > -m$ such that $f(k) = k + \delta$. The case $\delta = 0$ is studied extensively in the work of Bollobas and Riordan.

Preferential attachment networks

The following preferential attachment network with fixed outdegree is studied in the work of Hooghiemstra, van der Hofstad et al. and uses parameters $\delta > -m$ where $m \geq 2$ is an integer.

Preferential attachment networks

The following preferential attachment network with fixed outdegree is studied in the work of Hooghiemstra, van der Hofstad et al. and uses parameters $\delta > -m$ where $m \geq 2$ is an integer.

- $\mathcal{G}_1$ consists of a single vertex with m self loops.
- In each further step, given $\mathcal{G}_N$, we insert one new vertex and then successively insert m edges connecting the new vertex to vertex $n \leq N$ with probability

$$\sim \ (\text{degree of vertex } n) + \delta$$

or to itself with probability

$$\sim \ (\text{current degree}) + \frac{\delta}{m}.$$

Preferential attachment networks

Theorem

Denoting by

$$p_k = \left(2 + \tfrac{\delta}{m}\right) \frac{\Gamma(k+\delta)\Gamma(m+2+\delta+\frac{\delta}{m})}{\Gamma(m+\delta)\Gamma(k+3+\delta+\frac{\delta}{m})} \text{ for } k \geq m$$

we have

$$\lim_{N\uparrow\infty} X_N(k) = p_k \quad \text{for all } k, \text{ in probability.}$$

In particular, the network is scale-free with power-law exponent

$$\tau = 3 + \frac{\delta}{m}.$$

This was first proved for $\delta = 0$ by Bollobas, Riordan, Spencer and Tusnady.

Preferential attachment networks

Our focus in this course will be on a preferential attachment network with variable outdegree introduced by Dereich and Mörters. The variability of the outdegree will be used to maximise independence in the network. This makes the model easier to study than the model with fixed outdegree. An immediate advantage is that nonlinear attachment rules can be handled.

Preferential attachment networks

Our focus in this course will be on a preferential attachment network with variable outdegree introduced by Dereich and Mörters.

We fix a concave function $f\colon \mathbb{N}\cup\{0\}\to(0,\infty)$ with $f(0)\le 1$ and

$$\Delta f(k) := f(k+1) - f(k) < 1 \qquad \text{for all } k\in\mathbb{N}.$$

At time $N=1$, we have a single vertex (labeled 1). In each time step $N\to N+1$ we

- add a new vertex labeled $N+1$, and
- for each $n\le N$ independently introduce an oriented edge from the new vertex $N+1$ to the old vertex n with probability

$$\frac{f(\text{indegree of } n \text{ at time } N)}{N}.$$

Preferential attachment networks

Our focus in this course will be on a preferential attachment network with variable outdegree introduced by Dereich and Mörters.

We fix a concave function $f\colon \mathbb{N} \cup \{0\} \to (0, \infty)$ with $f(0) \leq 1$ and

$$\Delta f(k) := f(k+1) - f(k) < 1 \qquad \text{for all } k \in \mathbb{N}.$$

At time $N = 1$, we have a single vertex (labeled 1). In each time step $N \to N+1$ we

- add a new vertex labeled $N+1$, and
- for each $n \leq N$ independently introduce an oriented edge from the new vertex $N+1$ to the old vertex n with probability

$$\frac{f(\text{indegree of } n \text{ at time } N)}{N}.$$

Example:

Preferential attachment networks

Our focus in this course will be on a preferential attachment network with variable outdegree introduced by Dereich and Mörters.

We fix a concave function $f\colon \mathbb{N}\cup\{0\} \to (0,\infty)$ with $f(0) \le 1$ and

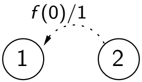
$$\Delta f(k) := f(k+1) - f(k) < 1 \qquad \text{for all } k \in \mathbb{N}.$$

At time $N = 1$, we have a single vertex (labeled 1). In each time step $N \to N+1$ we

- add a new vertex labeled $N+1$, and
- for each $n \le N$ independently introduce an oriented edge from the new vertex $N+1$ to the old vertex n with probability

$$\frac{f(\text{indegree of } n \text{ at time } N)}{N}.$$

Example:



Preferential attachment networks

Our focus in this course will be on a preferential attachment network with variable outdegree introduced by Dereich and Mörters.

We fix a concave function $f\colon \mathbb{N}\cup\{0\}\to(0,\infty)$ with $f(0)\le 1$ and

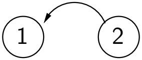
$$\Delta f(k) := f(k+1) - f(k) < 1 \qquad \text{for all } k \in \mathbb{N}.$$

At time $N=1$, we have a single vertex (labeled 1). In each time step $N\to N+1$ we

- add a new vertex labeled $N+1$, and
- for each $n\le N$ independently introduce an oriented edge from the new vertex $N+1$ to the old vertex n with probability

$$\frac{f(\text{indegree of } n \text{ at time } N)}{N}.$$

Example:



Preferential attachment networks

Our focus in this course will be on a preferential attachment network with variable outdegree introduced by Dereich and Mörters.

We fix a concave function $f : \mathbb{N} \cup \{0\} \to (0, \infty)$ with $f(0) \leq 1$ and

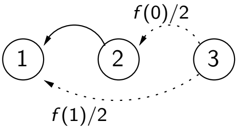
$$\Delta f(k) := f(k+1) - f(k) < 1 \quad \text{for all } k \in \mathbb{N}.$$

At time $N = 1$, we have a single vertex (labeled 1). In each time step $N \to N+1$ we

- add a new vertex labeled $N+1$, and
- for each $n \leq N$ independently introduce an oriented edge from the new vertex $N+1$ to the old vertex n with probability

$$\frac{f(\text{indegree of } n \text{ at time } N)}{N}.$$

Example:



Preferential attachment networks

Our focus in this course will be on a preferential attachment network with variable outdegree introduced by Dereich and Mörters.

We fix a concave function $f\colon \mathbb{N}\cup\{0\}\to(0,\infty)$ with $f(0)\le 1$ and

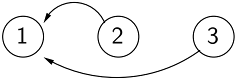
$$\Delta f(k) := f(k+1) - f(k) < 1 \qquad \text{for all } k\in\mathbb{N}.$$

At time $N=1$, we have a single vertex (labeled 1). In each time step $N\to N+1$ we

- add a new vertex labeled $N+1$, and
- for each $n\le N$ independently introduce an oriented edge from the new vertex $N+1$ to the old vertex n with probability

$$\frac{f(\text{indegree of } n \text{ at time } N)}{N}.$$

Example:



Preferential attachment networks

Our focus in this course will be on a preferential attachment network with variable outdegree introduced by Dereich and Mörters.

We fix a concave function $f \colon \mathbb{N} \cup \{0\} \to (0, \infty)$ with $f(0) \leq 1$ and

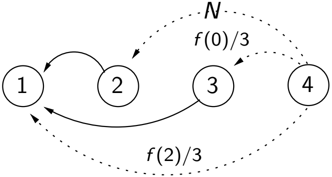
$$\Delta f(k) := f(k+1) - f(k) < 1 \qquad \text{for all } k \in \mathbb{N}.$$

At time $N = 1$, we have a single vertex (labeled 1). In each time step $N \to N+1$ we

- add a new vertex labeled $N+1$, and
- for each $n \leq N$ independently introduce an oriented edge from the new vertex $N+1$ to the old vertex n with probability

$$\frac{f(\text{indegree of } n \text{ at time } N)}{N}.$$

Example:



Preferential attachment networks

Our focus in this course will be on a preferential attachment network with variable outdegree introduced by Dereich and Mörters.

We fix a concave function $f \colon \mathbb{N} \cup \{0\} \to (0, \infty)$ with $f(0) \leq 1$ and

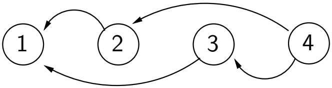
$$\Delta f(k) := f(k+1) - f(k) < 1 \quad \text{for all } k \in \mathbb{N}.$$

At time $N = 1$, we have a single vertex (labeled 1). In each time step $N \to N+1$ we

- add a new vertex labeled $N+1$, and
- for each $n \leq N$ independently introduce an oriented edge from the new vertex $N+1$ to the old vertex n with probability

$$\frac{f(\text{indegree of } n \text{ at time } N)}{N}.$$

Example:



Preferential attachment networks

The empirical indegree distribution of $\mathcal{G}_N$ is given by

$$X_N^{\mathrm{in}}(k) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{\text{indegree of vertex } i = k\}.$$

The empirical indegree distribution of $\mathcal{G}_N$ is given by

$$X_N^{\text{in}}(k) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{\text{indegree of vertex } i = k\}.$$

Theorem 1

Denoting by

$$\mu(k) = \frac{1}{1+f(k)} \prod_{l=0}^{k-1} \frac{f(l)}{1+f(l)},$$

we have

$$\lim_{N\uparrow\infty} X_N^{\text{in}}(k) = \mu(k) \qquad \text{for all } k, \text{ in probability.}$$

Preferential attachment networks

The empirical indegree distribution of $\mathcal{G}_N$ is given by

$$X_N^{in}(k) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{\text{indegree of vertex } i = k\}.$$

Theorem 1

Denoting by

$$\mu(k) = \frac{1}{1+f(k)} \prod_{l=0}^{k-1} \frac{f(l)}{1+f(l)},$$

we have

$$\lim_{N\uparrow\infty} X_N^{in}(k) = \mu(k) \quad \text{for all } k, \text{ in probability.}$$

The limit

$$\gamma := \lim_{k\uparrow\infty} \frac{f(k)}{k} = \inf_{n\geq 1} \Delta f(n)$$

exists by concavity and, by Theorem 1, under the assumption that $\gamma > 0$ the network is scale-free with power-law exponent

$$\tau = \frac{1+\gamma}{\gamma}.$$

Preferential attachment networks

The empirical indegree distribution of $\mathcal{G}_N$ is given by

$$X_N^{in}(k) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}\{\text{indegree of vertex } i = k\}.$$

Theorem 1

Denoting by

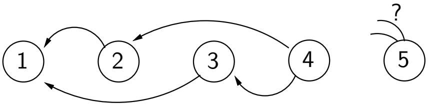
$$\mu(k) = \frac{1}{1+f(k)} \prod_{l=0}^{k-1} \frac{f(l)}{1+f(l)},$$

we have

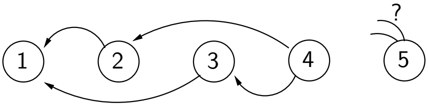
$$\lim_{N\uparrow\infty} X_N^{in}(k) = \mu(k) \qquad \text{for all } k, \text{ in probability.}$$

If $f(k) \sim k^{\alpha}$ for $0 \leq \alpha < 1$, then $\log \mu(k) \sim -\frac{1}{1-\alpha}\, k^{1-\alpha}$.

Preferential attachment networks



Preferential attachment networks

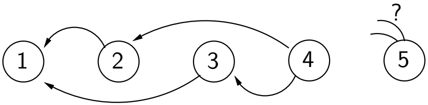


Theorem 2

The conditional distribution of the outdegree of the vertes with label $N + 1$, given the graph at time N, converges almost surely in the variational topology to the Poisson distribution with parameter

$$\sum_{k=0}^{\infty} \mu(k) f(k).$$

Preferential attachment networks



Theorem 2

The conditional distribution of the outdegree of the vertes with label $N+1$, given the graph at time N, converges almost surely in the variational topology to the Poisson distribution with parameter

$$\sum_{k=0}^{\infty} \mu(k) f(k).$$

The outdegree distribution is therefore light-tailed and does not interfere with the power-law exponent.

Preferential attachment networks

A small selection of references:

- Barabasi, Albert. Emergence of scaling in random networks. Science 286, 509–512 (1999)
- Bollobas, Riordan, Spencer, Tusnady. The degree sequence of a scale-free random graph process. Random Structures Algorithms 18, 279–290 (2001).
- Dereich, Mörters. Random networks with sublinear preferential attachment: Degree evolutions. Electronic Journal of Probability 14, 1222-1267 (2009).
- Krapivsky, Redner. Organization of growing random networks. Phys. Rev. E 63, 066123 (2001).