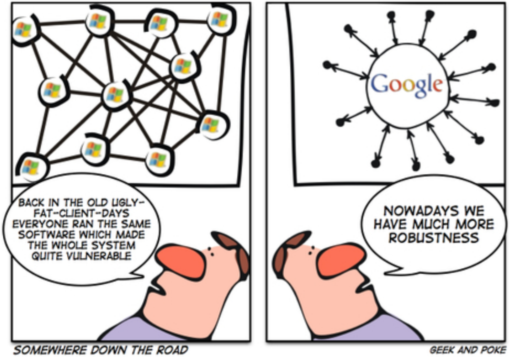
Structure and dynamics of complex networks

April 7, 2020

ROBUSTNESS AND PERCOLATION



WHAT IS ROBUSTNESS?

Robustness of complex networks

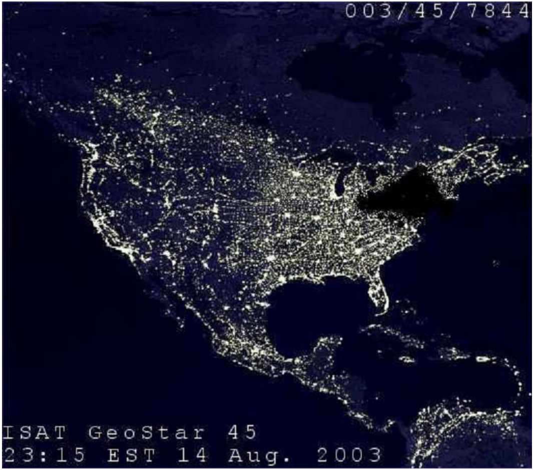
Some complex networks seem to be robust in the following general sense:

- Cell: gene mutations...
- Internet: router breakdowns...
- etc.

Although a part of the nodes (links) are down, the system as a whole is functioning normally.

Robustness of complex networks

However, if too many nodes or links are removed, the whole system breaks down:



Robustness of complex networks

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

Robustness of complex networks

- What does it mean that a network is "still functioning", and what does it mean that a network is "down"?

Robustness of complex networks

- What does it mean that a network is "still functioning", and what does it mean that a network is "down"?

→ If we are interested in the details, these can have a different meaning in each different system...

Robustness of complex networks

- What does it mean that a network is "still functioning", and what does it mean that a network is "down"?

→ If we are interested in the details, these can have a different meaning in each different system...

→ However, in general probably the most fundamental "function" or "role" of a network is to provide access for its nodes to each other.

Robustness of complex networks

- What does it mean that a network is "still functioning", and what does it mean that a network is "down"?

→ If we are interested in the details, these can have a different meaning in each different system...

→ However, in general probably the most fundamental "function" or "role" of a network is to provide access for its nodes to each other.

Thus, as long as the network contains a **giant component**, it is functioning, since the existence of the giant component ensures that most of the nodes can reach each other.

Robustness of complex networks

- What does it mean that a network is "still functioning", and what does it mean that a network is "down"?

→ If we are interested in the details, these can have a different meaning in each different system...

→ However, in general probably the most fundamental "function" or "role" of a network is to provide access for its nodes to each other.

Thus, as long as the network contains a **giant component**, it is functioning, since the existence of the giant component ensures that most of the nodes can reach each other.

→ Based on this we can give a simple formulation of what we mean by the robustness of a complex network.

Robustness of complex networks

- When discussing network robustness, we imagine a **random node removal process**, imitating the subsequent break down or malfunctioning of the nodes.

Robustness of complex networks

Robustness of complex networks

- When discussing network robustness, we imagine a **random node removal process**, imitating the subsequent break down or malfunctioning of the nodes.
- In this process we take out the nodes one by one, always choosing the next node uniformly at random. Let us denote the fraction of removed nodes by f, where $f = 0$ at the beginning of the removal process, and f is increasing as we carry on with the subsequent removals, reaching $f = 1$ if we take out all nodes in the system.

Robustness of complex networks

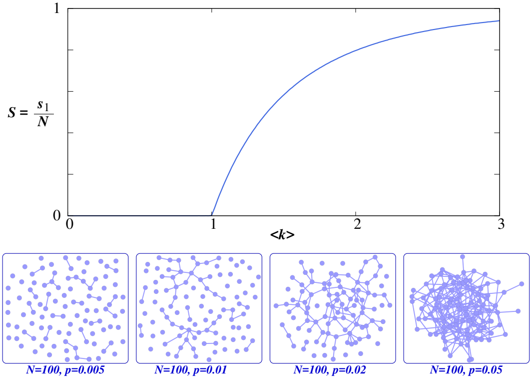
- When discussing network robustness, we imagine a **random node removal process**, imitating the subsequent break down or malfunctioning of the nodes.
- In this process we take out the nodes one by one, always choosing the next node uniformly at random. Let us denote the fraction of removed nodes by f, where $f = 0$ at the beginning of the removal process, and f is increasing as we carry on with the subsequent removals, reaching $f = 1$ if we take out all nodes in the system.
- As long as the network contains a **giant component**, it is considered to be still functioning, however, when the giant component is destroyed, the system is down.

Robustness of complex networks

- When discussing network robustness, we imagine a **random node removal process**, imitating the subsequent break down or malfunctioning of the nodes.
- In this process we take out the nodes one by one, always choosing the next node uniformly at random. Let us denote the fraction of removed nodes by f, where $f = 0$ at the beginning of the removal process, and f is increasing as we carry on with the subsequent removals, reaching $f = 1$ if we take out all nodes in the system.
- As long as the network contains a **giant component**, it is considered to be still functioning, however, when the giant component is destroyed, the system is down.
- Thus, the **critical f_c at which the giant component is destroyed** provides a simple indicator of the robustness:
  - If f_c is small, the network is fragile,
  - If f_c is large, the network is robust.

Robustness of complex networks

Since the giant component is a really important concept in this chapter, before actually examining the robustness of complex networks, it is a good idea to deepen our knowledge about the emergence or disappearance of the giant component in networks.



GIANT COMPONENT IN THE ERDŐS-RÉNYI GRAPH

Giant component in the E-R graph

- Does the Erdős-Rényi graph contain a giant component?

Giant component in the E-R graph

- Does the Erdős-Rényi graph contain a giant component?

- In the two limiting cases its easy to answer this question:
  - if the linking probability p is $p = 0$, all nodes are isolated, thus, we have no giant component.
  - if the linking probability p is $p = 1$, all nodes are connected to each other, thus, the whole network is a single giant component.

Giant component in the E-R graph

- Does the Erdős-Rényi graph contain a giant component?

- In the two limiting cases its easy to answer this question:
  - if the linking probability p is $p = 0$, all nodes are isolated, thus, we have no giant component.
  - if the linking probability p is $p = 1$, all nodes are connected to each other, thus, the whole network is a single giant component.

- What about in between the limiting cases?

Giant component in the E-R graph

- Does the Erdős-Rényi graph contain a giant component?

- In the two limiting cases its easy to answer this question:
  - if the linking probability p is $p = 0$, all nodes are isolated, thus, we have no giant component.
  - if the linking probability p is $p = 1$, all nodes are connected to each other, thus, the whole network is a single giant component.

- What about in between the limiting cases?

$\rightarrow$ Let's find out!

Giant component in the E-R graph

First two technical remarks:

Giant component in the E-R graph

First two technical remarks:

- It turns out that in the E-R model, it is better to examine the components as a function of the average degree $\langle k \rangle$ and not as a function of p.

  Although for any given E-R graph size N we have a simple linear relation between p and $\langle k \rangle$ in the form of $\langle k \rangle = Np$, when comparing E-R graphs of different size, it is more natural to take $\langle k \rangle$ as the other parameter beside N instead of p.

First two technical remarks:

- It turns out that in the E-R model, it is better to examine the components as a function of the average degree $\langle k \rangle$ and not as a function of p.

  Although for any given E-R graph size N we have a simple linear relation between p and $\langle k \rangle$ in the form of $\langle k \rangle = Np$, when comparing E-R graphs of different size, it is more natural to take $\langle k \rangle$ as the other parameter beside N instead of p.

- Similarly, when measuring the sizes of the components, it is a good idea to switch for any component α from the number of nodes in the component s_α to the relative component size, given by

$$S_\alpha = \frac{s_\alpha}{N}.$$

  Again, when comparing E-R graphs of different sizes this is a natural thing to do, since the relative size is always between 0 and 1.

Giant component in the E-R graph

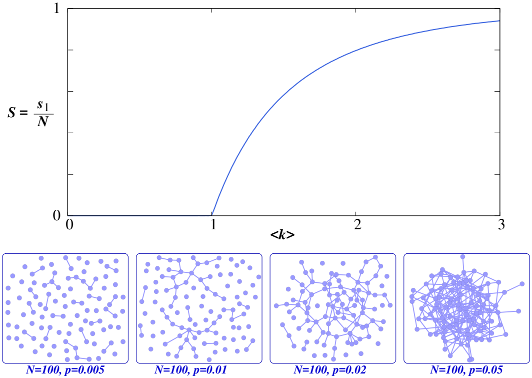
Relative size of the **largest component** as a function of $\langle k \rangle$:

- Please take a look at the accompanying notebook where we examine this by simulations for finite size.

Giant component in the E-R graph

Relative size of the **largest component** as a function of $\langle k \rangle$:

- Please take a look at the accompanying notebook where we examine this by simulations for finite size.
- In case of an infinitely large E-R graph we would obtain the following curve (where the graphs below are just illustrations):



Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:

    - it is either non existent

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:

    - it is either non existent $\rightarrow$ probability$= 1 - p$,

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:
    - it is either non existent $\rightarrow$ probability$= 1 - p$,
    - or j is also not in the giant component

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:
    - it is either non existent → probability$= 1 - p$,
    - or j is also not in the giant component → probability$= pu$.

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:
    - it is either non existent $\rightarrow$ probability$= 1 - p$,
    - or j is also not in the giant component $\rightarrow$ probability$= pu$.

$\rightarrow$ The probability that i is NOT in the giant component is

$$(1 - p + pu)^{N-1}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:
    - it is either non existent → probability$= 1 - p$,
    - or j is also not in the giant component → probability$= pu$.

→ The probability that i is NOT in the giant component is

$$(1 - p + pu)^{N-1}$$

  - However, the probability that i is NOT in the giant component is u by definition, thus

$$u = (1 - p + pu)^{N-1}.$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$u = (1 - p(1-u))^{N-1} = \left(1 - \frac{\langle k \rangle}{N-1}(1-u)\right)^{N-1}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$\begin{aligned} u &= (1 - p(1-u))^{N-1} = \left(1 - \frac{\langle k \rangle}{N-1}(1-u)\right)^{N-1} \\ \ln u &= (N-1)\ln\left[1 - \frac{\langle k \rangle}{N-1}(1-u)\right] \end{aligned}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
    - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$\begin{aligned}
u &= (1 - p(1-u))^{N-1} = \left(1 - \frac{\langle k \rangle}{N-1}(1-u)\right)^{N-1} \\
\ln u &= (N-1)\ln\left[1 - \frac{\langle k \rangle}{N-1}(1-u)\right] \\
\ln u &\approx -(N-1)\left[\frac{\langle k \rangle}{N-1}(1-u)\right]
\end{aligned}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$\begin{aligned}
u &= (1 - p(1-u))^{N-1} = \left(1 - \frac{\langle k \rangle}{N-1}(1-u)\right)^{N-1} \\
\ln u &= (N-1)\ln\left[1 - \frac{\langle k \rangle}{N-1}(1-u)\right] \\
\ln u &\approx -(N-1)\left[\frac{\langle k \rangle}{N-1}(1-u)\right] \\
u &\approx e^{-\langle k \rangle (1-u)}
\end{aligned}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$u \approx e^{-\langle k \rangle (1-u)}.$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$u \approx e^{-\langle k \rangle (1-u)}.$$

$$S \approx 1 - e^{-\langle k \rangle S}.$$

Giant component in the E-R graph

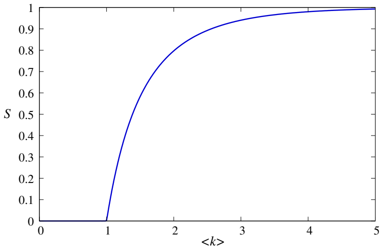
- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$u \approx e^{-\langle k \rangle (1-u)}.$$

$$S \approx 1 - e^{-\langle k \rangle S}.$$

This equation can be solved numerically for any $\langle k \rangle$, giving the S in the figure.
Please take a look at the jupyter notebook, where we calculate S.



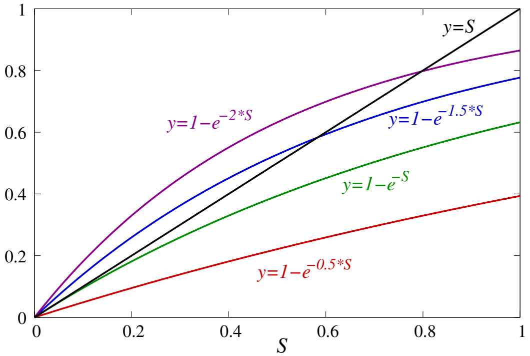
- **Where is the transition point?**

Giant component in the E-R graph

- **Where is the transition point?**
  - The transition point is where S becomes larger than zero.

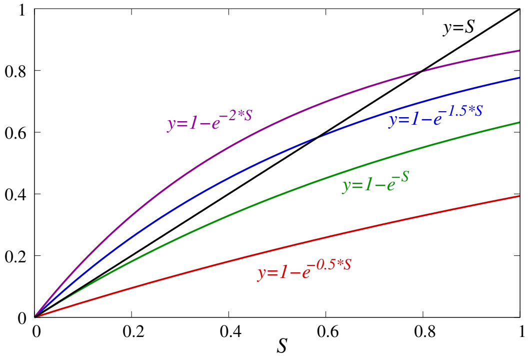
Giant component in the E-R graph

- **Where is the transition point?**
  - The transition point is where S becomes larger than zero.
  - Let us solve $S = 1 - e^{-\langle k \rangle S}$ "graphically":



Giant component in the E-R graph

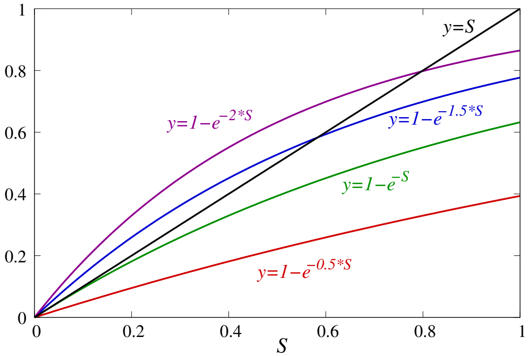
- **Where is the transition point?**
  - The transition point is where S becomes larger than zero.
  - Let us solve $S = 1 - e^{-\langle k \rangle S}$ "graphically":



→ What is the condition for having a non-trivial solution?

Giant component in the E-R graph

- **Where is the transition point?**
  - The transition point is where S becomes larger than zero.
  - Let us solve $S = 1 - e^{-\langle k \rangle S}$ "graphically":



→ What is the condition for having a non-trivial solution?

$$\frac{d}{dS}\left(1 - e^{-\langle k \rangle S}\right) \geq 1 \Big|_{S=0}$$

Giant component in the E-R graph

- **Where is the transition point?**
  - The transition point is where S becomes larger than zero.

$$\frac{d}{dS}\left(1 - e^{-\langle k \rangle S}\right) \geq 1 \Big|_{S=0}$$

Giant component in the E-R graph

- **Where is the transition point?**
  - The transition point is where S becomes larger than zero.

$$\frac{d}{dS}\left(1-e^{-\langle k\rangle S}\right) \geq 1\Big|_{S=0}$$

$$\langle k\rangle\, e^{-\langle k\rangle S}\Big|_{S=0} \geq 1$$

Giant component in the E-R graph

- **Where is the transition point?**
  - The transition point is where S becomes larger than zero.

$$\frac{d}{dS}\left(1 - e^{-\langle k \rangle S}\right) \geq 1 \Big|_{S=0}$$

$$\langle k \rangle \, e^{-\langle k \rangle S}\Big|_{S=0} \geq 1$$

$$\rightarrow \langle k \rangle \geq 1$$

Giant component in the E-R graph

- **Where is the transition point?**
  - The transition point is where S becomes larger than zero.

$$\frac{d}{dS}\left(1 - e^{-\langle k \rangle S}\right) \geq 1 \Big|_{S=0}$$

$$\langle k \rangle \, e^{-\langle k \rangle S}\Big|_{S=0} \geq 1$$

$$\rightarrow \langle k \rangle \geq 1$$

- **Thus, the transition point is at $k = 1$, and for $k \geq 1$ we have a giant component in the E-R graph.**

PERCOLATION TRANSITION

Phase transitions

- For a physicist, the emergence of the giant component in the E-R model at $\langle k \rangle = 1$ is like a **phase transition**!

Phase transitions

- For a physicist, the emergence of the giant component in the E-R model at $\langle k \rangle = 1$ is like a **phase transition**!

→ What are phase transitions?

Phase transitions

- For a physicist, the emergence of the giant component in the E-R model at $\langle k \rangle = 1$ is like a **phase transition**!

$\rightarrow$ What are phase transitions?

- E.g, the freezing of water (the two phases are liquid and solid),

Phase transitions

- For a physicist, the emergence of the giant component in the E-R model at $\langle k \rangle = 1$ is like a **phase transition**!

→ What are phase transitions?

  - E.g, the freezing of water (the two phases are liquid and solid),
  - or the "freezing" of the traffic on a highway in case of a traffic jam (the two phases are the jammed traffic and the fast flowing traffic),

Phase transitions

- For a physicist, the emergence of the giant component in the E-R model at $\langle k \rangle = 1$ is like a **phase transition**!

→ What are phase transitions?

  - E.g, the freezing of water (the two phases are liquid and solid),
  - or the "freezing" of the traffic on a highway in case of a traffic jam (the two phases are the jammed traffic and the fast flowing traffic),
  - or ferromagnetic transition when we heat a magnet: above a certain temperature T_c it looses its magnetisation and becomes similar to a non magnetic material (the two phases are magnetic and non magnetic),
  - etc.

Phase transitions

- For a physicist, the emergence of the giant component in the E-R model at $\langle k \rangle = 1$ is like a **phase transition**!

→ What are phase transitions?

  - E.g, the freezing of water (the two phases are liquid and solid),
  - or the "freezing" of the traffic on a highway in case of a traffic jam (the two phases are the jammed traffic and the fast flowing traffic),
  - or ferromagnetic transition when we heat a magnet: above a certain temperature T_c it looses its magnetisation and becomes similar to a non magnetic material (the two phases are magnetic and non magnetic),
  - etc.

- Phase transitions have several **universal features**, and we can discuss them in a more or less uniform framework.

Phase transitions

- Every phase transition has a **control parameter**, and by changing this parameter we can move from one phase to the other:

Phase transitions

- Every phase transition has a **control parameter**, and by changing this parameter we can move from one phase to the other:
  - in case of freezing of the water and the ferromagnetic transition this is the temperature,

Phase transitions

- Every phase transition has a **control parameter**, and by changing this parameter we can move from one phase to the other:
  - in case of freezing of the water and the ferromagnetic transition this is the temperature,
  - in case of the traffic jam it is the density of the vehicles,

Phase transitions

- Every phase transition has a **control parameter**, and by changing this parameter we can move from one phase to the other:
  - in case of freezing of the water and the ferromagnetic transition this is the temperature,
  - in case of the traffic jam it is the density of the vehicles,
  - in case of the E-R graph this is the average degree, $\langle k \rangle$.

Phase transitions

- For every phase transition we can define an **order parameter**, which is like an indicator variable that has a value of 0 in one phase, and value larger than 0 in the other phase:

Phase transitions

- For every phase transition we can define an **order parameter**, which is like an indicator variable that has a value of 0 in one phase, and value larger than 0 in the other phase:
  - in case of the ferromagnetic transition this can be the magnetism (which is 0 above T_c and non-zero below T_c),

Phase transitions

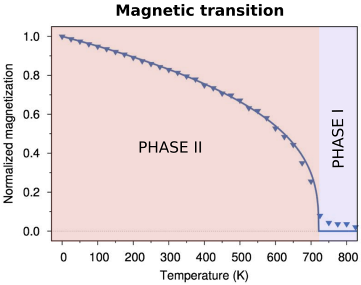
- For every phase transition we can define an **order parameter**, which is like an indicator variable that has a value of 0 in one phase, and value larger than 0 in the other phase:
  - in case of the ferromagnetic transition this can be the magnetism (which is 0 above T_c and non-zero below T_c),
  - in case of the traffic jam this can be the size of the jammed region,

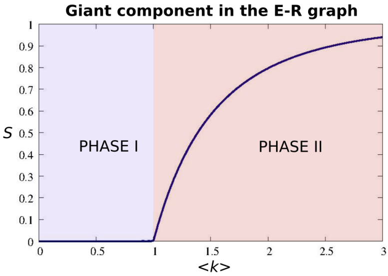
Phase transitions

- For every phase transition we can define an **order parameter**, which is like an indicator variable that has a value of 0 in one phase, and value larger than 0 in the other phase:
  - in case of the ferromagnetic transition this can be the magnetism (which is 0 above T_c and non-zero below T_c),
  - in case of the traffic jam this can be the size of the jammed region,
  - in case of the E-R graph this can be the (relative) size of the largest connected component.

Phase transitions

When plotting the **order parameter** as a function of the **control parameter**, phase transitions look quite similar:





Percolation transition

Beside the ferromagnetic transition, there is another phase transition in statistical physics that is even more closely related to the giant component in the E-R graph:

THE PERCOLATION TRANSITION!!

What is percolation?

Percolation

What is percolation?

- Percolation is the passing of a liquid through some porous material or filter.

What is percolation?

- Percolation is the passing of a liquid through some porous material or filter.

→ Thus, when making coffee, you are actually percolating water through ground coffee beans, and this device for making coffee (often found in Hungarian households) is called a percolator:

Percolation

Although coffee is important enough by itself for physicists, percolation has applications also in industry:

Percolation

Although coffee is important enough by itself for physicists, percolation has applications also in industry:

- If oil is found in porous rock, a simple idea is to apply pressure (making the oil percolate through the pores) and pump it out.

Percolation

Although coffee is important enough by itself for physicists, percolation has applications also in industry:

- If oil is found in porous rock, a simple idea is to apply pressure (making the oil percolate through the pores) and pump it out.
- However, this works only if the small pores are connected enough so that together they make up a large connected cavity...

Percolation

Although coffee is important enough by itself for physicists, percolation has applications also in industry:

- If oil is found in porous rock, a simple idea is to apply pressure (making the oil percolate through the pores) and pump it out.
- However, this works only if the small pores are connected enough so that together they make up a large connected cavity...
- So before actually setting up the machinery for oil mining in this way, it is good to know whether this is the case or not. Actually it turns out that by examining the structure of a smaller rock sample we can make a pretty good guess about that.

Percolation

Although coffee is important enough by itself for physicists, percolation has applications also in industry:

- If oil is found in porous rock, a simple idea is to apply pressure (making the oil percolate through the pores) and pump it out.
- However, this works only if the small pores are connected enough so that together they make up a large connected cavity...
- So before actually setting up the machinery for oil mining in this way, it is good to know whether this is the case or not. Actually it turns out that by examining the structure of a smaller rock sample we can make a pretty good guess about that.
- The theoretical models for the large scale structures made up by small cavities in porous materials are called **percolation models**.

Percolation model

- The simplest idea is to model the structure of a porous material with a **lattice** in an Euclidean space (e.g., a 3d lattice), where the lattice points are the possible places for having small cavities.

Percolation model

- The simplest idea is to model the structure of a porous material with a **lattice** in an Euclidean space (e.g., a 3d lattice), where the lattice points are the possible places for having small cavities.
- We introduce a probability p for the lattice points to be occupied by a cavity, and the **cavities are placed absolutely at random**.

Percolation model

- The simplest idea is to model the structure of a porous material with a **lattice** in an Euclidean space (e.g., a 3d lattice), where the lattice points are the possible places for having small cavities.
- We introduce a probability p for the lattice points to be occupied by a cavity, and the **cavities are placed absolutely at random**.

  Thus, at $p = 0$ we have no cavities at all, whereas at $p = 1$ all lattice points are occupied by a small cavity.

Percolation model

- The simplest idea is to model the structure of a porous material with a **lattice** in an Euclidean space (e.g., a 3d lattice), where the lattice points are the possible places for having small cavities.
- We introduce a probability p for the lattice points to be occupied by a cavity, and the **cavities are placed absolutely at random**.

  Thus, at $p = 0$ we have no cavities at all, whereas at $p = 1$ all lattice points are occupied by a small cavity.
- **Adjacent cavities are assumed to be connected to each other**. Thus, for large p the small cavities together make up a **large connected cavity, also called as cluster, that is spanning from one side of the lattice to the other**, making it possible to percolate a liquid through the lattice.

Percolation model

- The simplest idea is to model the structure of a porous material with a **lattice** in an Euclidean space (e.g., a 3d lattice), where the lattice points are the possible places for having small cavities.
- We introduce a probability p for the lattice points to be occupied by a cavity, and the **cavities are placed absolutely at random**.

  Thus, at $p = 0$ we have no cavities at all, whereas at $p = 1$ all lattice points are occupied by a small cavity.
- **Adjacent cavities are assumed to be connected to each other**. Thus, for large p the small cavities together make up a **large connected cavity, also called as cluster, that is spanning from one side of the lattice to the other**, making it possible to percolate a liquid through the lattice.
- However, when p is small, we are likely to have only small isolated clusters of cavities, and we cannot percolate.

Percolation model

- The simplest idea is to model the structure of a porous material with a **lattice** in an Euclidean space (e.g., a 3d lattice), where the lattice points are the possible places for having small cavities.
- We introduce a probability p for the lattice points to be occupied by a cavity, and the **cavities are placed absolutely at random**.

  Thus, at $p = 0$ we have no cavities at all, whereas at $p = 1$ all lattice points are occupied by a small cavity.
- **Adjacent cavities are assumed to be connected to each other**. Thus, for large p the small cavities together make up a **large connected cavity, also called as cluster, that is spanning from one side of the lattice to the other**, making it possible to percolate a liquid through the lattice.
- However, when p is small, we are likely to have only small isolated clusters of cavities, and we cannot percolate.

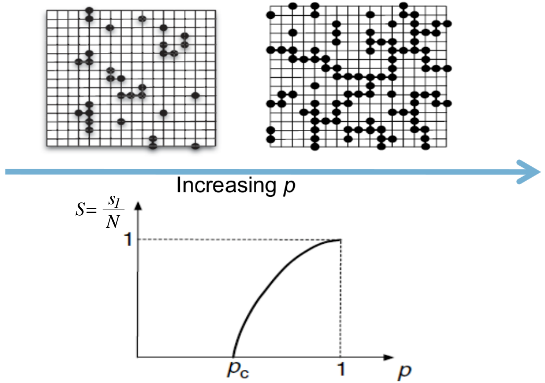
→ We have **two phases** (percolating cluster and small clusters), and the transition from one to the other is called

THE PERCOLATION TRANSITION!

Percolation transition

Percolation on lattices

- a node is occupied with probability p,
- at p_c a spanning cluster emerges.
- control parameter is p, the order parameter is the relative size of the largest connected cluster.



Percolation transition

Analogy between the percolation transition and the emergence of the giant component in the E-R graph:

	Percolation		**E-R graph**
phase I.:	no giant cluster	$\leftrightarrow$	no giant component
phase II.:	giant cluster	$\leftrightarrow$	giant component
control param.:	p, site occup. prob.	$\leftrightarrow$	$\langle k\rangle$, average degree
order param.:	rel. size of giant cluster	$\leftrightarrow$	rel. size of giant comp.

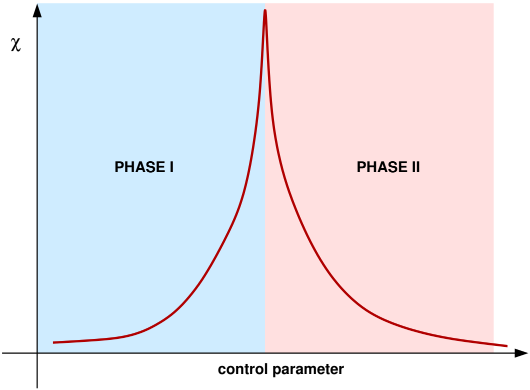
Percolation transition

Analogy between the percolation transition and the emergence of the giant component in the E-R graph:

	Percolation		**E-R graph**
phase I.:	no giant cluster	$\leftrightarrow$	no giant component
phase II.:	giant cluster	$\leftrightarrow$	giant component
control param.:	p, site occup. prob.	$\leftrightarrow$	$\langle k\rangle$, average degree
order param.:	rel. size of giant cluster	$\leftrightarrow$	rel. size of giant comp.

Although no one ever tried to percolate a liquid through a network, still, based on this analogy, the emergence of the giant component in the E-R graph at $\langle k\rangle = 1$ is often referred to as

THE PERCOLATION TRANSITION OF THE E-R GRAPH !



SUSCEPTIBILITY

Susceptibility

One of the surprising facts about phase transitions is that close to the transition point the system can show **extreme large sensitivity** with respect to outside **perturbations that drive the system towards the other phase**:

Susceptibility

One of the surprising facts about phase transitions is that close to the transition point the system can show **extreme large sensitivity** with respect to outside **perturbations that drive the system towards the other phase**:

- just a single flip of an "atomic magnet" in the crystal lattice can induce the build up of macroscopic magnetism at the critical temperature of a ferromagnet,

Susceptibility

One of the surprising facts about phase transitions is that close to the transition point the system can show **extreme large sensitivity** with respect to outside **perturbations that drive the system towards the other phase**:

- just a single flip of an "atomic magnet" in the crystal lattice can induce the build up of macroscopic magnetism at the critical temperature of a ferromagnet,
- insertion of just a single extra car can induce a large traffic jam,

Susceptibility

One of the surprising facts about phase transitions is that close to the transition point the system can show **extreme large sensitivity** with respect to outside **perturbations that drive the system towards the other phase**:

- just a single flip of an "atomic magnet" in the crystal lattice can induce the build up of macroscopic magnetism at the critical temperature of a ferromagnet,
- insertion of just a single extra car can induce a large traffic jam,
- insertion of just a single extra link can induce the emergence of the giant component in the random graph.

Susceptibility

- The quantity measuring the sensitivity of the system is called the **susceptibility**, usually denoted by χ.

Susceptibility

- The quantity measuring the sensitivity of the system is called the **susceptibility**, usually denoted by χ.
- Similarly to the control parameter and the order parameter, the actual definition or formula of the susceptibility is different for each system.

Susceptibility

- The quantity measuring the sensitivity of the system is called the **susceptibility**, usually denoted by χ.
- Similarly to the control parameter and the order parameter, the actual definition or formula of the susceptibility is different for each system.
- Nevertheless, the behaviour of the susceptibility is the same across all systems:

  When plotted as a function of the control parameter, it has a **strong, sharp peak at the transition point**.

  For macroscopic (or more precisely, infinitely large) systems χ is actually divergent at the transition point.

Susceptibility

- **How should we define the susceptibility for the percolation transition of the E-R model?**

Susceptibility

- **How should we define the susceptibility for the percolation transition of the E-R model?**

→ A perturbation that would drive a yet "dispersed" E-R graph towards the transition point would be something like this:

Susceptibility

- **How should we define the susceptibility for the percolation transition of the E-R model?**

→ A perturbation that would drive a yet "dispersed" E-R graph towards the transition point would be something like this:

  - choose a node at random, sitting in some component,

Susceptibility

- **How should we define the susceptibility for the percolation transition of the E-R model?**

→ A perturbation that would drive a yet "dispersed" E-R graph towards the transition point would be something like this:

  - choose a node at random, sitting in some component,
  - choose a 2^{nd} node at random, however make sure that it belongs to a different component,

Susceptibility

- **How should we define the susceptibility for the percolation transition of the E-R model?**

→ A perturbation that would drive a yet "dispersed" E-R graph towards the transition point would be something like this:

- choose a node at random, sitting in some component,
- choose a 2^{nd} node at random, however make sure that it belongs to a different component,
- connect these two nodes with a link, thereby merging the two involved components into a single (larger) component.

Susceptibility

- **How should we define the susceptibility for the percolation transition of the E-R model?**

→ A perturbation that would drive a yet "dispersed" E-R graph towards the transition point would be something like this:

  - choose a node at random, sitting in some component,
  - choose a 2^{nd} node at random, however make sure that it belongs to a different component,
  - connect these two nodes with a link, thereby merging the two involved components into a single (larger) component.

- How can we measure the "sensitivity" of the system with respect to such a perturbation?

Susceptibility

- **How should we define the susceptibility for the percolation transition of the E-R model?**

→ A perturbation that would drive a yet "dispersed" E-R graph towards the transition point would be something like this:

  - choose a node at random, sitting in some component,
  - choose a 2^{nd} node at random, however make sure that it belongs to a different component,
  - connect these two nodes with a link, thereby merging the two involved components into a single (larger) component.

- How can we measure the "sensitivity" of the system with respect to such a perturbation?

→ **A simple idea is to define χ as the expected change in the component size for the initially separated components that are merged because of the extra added link.**

- Based on the previous slide we can formulate χ as follows:

Susceptibility

- Based on the previous slide we can formulate χ as follows:
  - Assuming that the probability distribution for the component sizes is given by

    $$p(s) = \frac{N_s}{N_{\text{comp}}},$$

    where N_s gives the number of components with size s, and N_{comp} is the overall number of components,

Susceptibility

- Based on the previous slide we can formulate χ as follows:
  - Assuming that the probability distribution for the component sizes is given by
    $$p(s)=\frac{N_s}{N_{\rm comp}},$$
    where N_s gives the number of components with size s, and $N_{\rm comp}$ is the overall number of components,
  - the probability for choosing the 2$^{\rm nd}$ node from a component of size s can be written as $s\cdot p(s)$ (since choosing the node from a given component is proportional to the size of the component).

Susceptibility

- Based on the previous slide we can formulate χ as follows:
  - Assuming that the probability distribution for the component sizes is given by
  $$p(s) = \frac{N_s}{N_{\text{comp}}},$$
  where N_s gives the number of components with size s, and N_{comp} is the overall number of components,
  - the probability for choosing the 2nd node from a component of size s can be written as $s \cdot p(s)$ (since choosing the node from a given component is proportional to the size of the component).
  - However, when choosing the 2nd node from a component of size s, the component size change from the point of view of the first node is s, thus, if we take into account all possibilities we obtain
  $$\chi = \sum s \cdot s \cdot p(s) = \sum s^2 p(s) = \langle s^2 \rangle,$$
  which means that χ is equal to the second moment of the component size distribution!

Susceptibility

- A very important remark about the calculation of χ:

  The giant component should be left out of the formula for χ!

Susceptibility

- A very important remark about the calculation of χ:

  The giant component should be left out of the formula for χ!

- If the giant component is not removed from the formula of χ, due to the large size of the giant component and the s^2 factor in the formula of χ, we would get extreme large χ values everywhere in the phase with the giant component. However, what we want is a sharp peak in χ at the transition point.

Susceptibility

- A very important remark about the calculation of χ:

  The giant component should be left out of the formula for χ!

- If the giant component is not removed from the formula of χ, due to the large size of the giant component and the s^2 factor in the formula of χ, we would get extreme large χ values everywhere in the phase with the giant component. However, what we want is a sharp peak in χ at the transition point.

- Thus, the final formula for χ can be written as

$$\chi = \sum_{s \neq s_{\max}} s^2 p(s).$$

Susceptibility

- A very important remark about the calculation of χ:

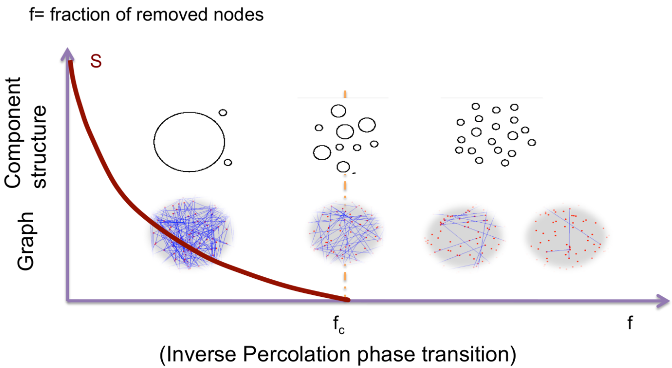
  The giant component should be left out of the formula for χ!

- If the giant component is not removed from the formula of χ, due to the large size of the giant component and the s^2 factor in the formula of χ, we would get extreme large χ values everywhere in the phase with the giant component. However, what we want is a sharp peak in χ at the transition point.

- Thus, the final formula for χ can be written as

$$\chi = \sum_{s \neq s_{\max}} s^2 p(s).$$

- Please take a look at the accompanying jupyter notebook, where we calculate and plot χ for the E-R graph.



RESILIENCE AGAINST RANDOM NODE REMOVAL

Robustness and resilience

Now that we have the necessary background knowledge about phase transitions, percolation, order and control parameters, susceptibility, etc., we can return to our main goal:

Robustness and resilience

Now that we have the necessary background knowledge about phase transitions, percolation, order and control parameters, susceptibility, etc., we can return to our main goal:

- Study the **robustness** of complex networks,

Robustness and resilience

Now that we have the necessary background knowledge about phase transitions, percolation, order and control parameters, susceptibility, etc., we can return to our main goal:

- Study the **robustness** of complex networks,
- where robustness is defined based on the the resilience against random node removal.

Robustness and resilience

Now that we have the necessary background knowledge about phase transitions, percolation, order and control parameters, susceptibility, etc., we can return to our main goal:

- Study the **robustness** of complex networks,
- where robustness is defined based on the resilience against random node removal.

→ In the light of percolation models, the random node removal process we defined at the beginning is basically an **"inverse percolation" transition**:

Robustness and resilience

Now that we have the necessary background knowledge about phase transitions, percolation, order and control parameters, susceptibility, etc., we can return to our main goal:

- Study the **robustness** of complex networks,
- where robustness is defined based on the resilience against random node removal.

→ In the light of percolation models, the random node removal process we defined at the beginning is basically an **"inverse percolation" transition**:

  - In the initial phase the network contains a giant component,

Robustness and resilience

Now that we have the necessary background knowledge about phase transitions, percolation, order and control parameters, susceptibility, etc., we can return to our main goal:

- Study the **robustness** of complex networks,
- where robustness is defined based on the resilience against random node removal.

→ In the light of percolation models, the random node removal process we defined at the beginning is basically an **"inverse percolation" transition**:
  - In the initial phase the network contains a giant component,
  - however, with subsequent random removal of the nodes our goal is to destroy this,

Robustness and resilience

Now that we have the necessary background knowledge about phase transitions, percolation, order and control parameters, susceptibility, etc., we can return to our main goal:

- Study the **robustness** of complex networks,
- where robustness is defined based on the resilience against random node removal.

→ In the light of percolation models, the random node removal process we defined at the beginning is basically an **"inverse percolation" transition**:

  - In the initial phase the network contains a giant component,
  - however, with subsequent random removal of the nodes our goal is to destroy this,
  - which means that we push the system towards the other phase, where instead of a giant component we have only isolated small components.

Robustness and resilience

Now that we have the necessary background knowledge about phase transitions, percolation, order and control parameters, susceptibility, etc., we can return to our main goal:

- Study the **robustness** of complex networks,
- where robustness is defined based on the resilience against random node removal.

→ In the light of percolation models, the random node removal process we defined at the beginning is basically an **"inverse percolation" transition**:

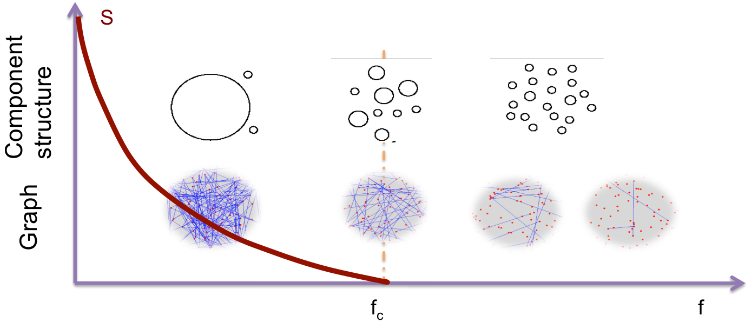
  - In the initial phase the network contains a giant component,
  - however, with subsequent random removal of the nodes our goal is to destroy this,
  - which means that we push the system towards the other phase, where instead of a giant component we have only isolated small components.

- **Please take a look at the accompanying jupyter notebook where we simulate the random node removal process on both a real network and an E-R graph.**

Random node removal: “inverse percolation"

Illustration:

f= fraction of removed nodes



(from the slides of A.-L. Barabási)

Condition for the existence of a giant component

- In the following we are going to derive a simple, general condition for the existence of a giant component in networks.
- What can we use that for?
- If we see that it is not fulfilled any more, then based on that we know that the giant component has been destroyed!

Condition for the existence of a giant component

- Assume a network with a giant component.
- The average degree in a giant component: $\langle k \rangle_{\mathrm{G}} \geq 2$

$\leftrightarrow$

- Let $\mathcal{P}(k_i \mid i \leftrightarrow j)$ be the **conditional probability** for node i to have degree k_i given it is connected to some random other node j.

- Suppose we have a **random network** with a **GC**. Assuming that a node connected to at least one other node is actually linked to the GC we obtain

$$\langle k \rangle_{\mathrm{G}} \simeq \langle k_i \mid i \leftrightarrow j \rangle = \sum_i k_i \mathcal{P}(k_i \mid i \leftrightarrow j) \geq 2.$$

Condition for the existence of a giant component

- Assume a network with a giant component.
- The average degree in a giant component: $\langle k \rangle_{\mathrm{G}} \geq 2$

$\leftrightarrow$

- Let $\mathcal{P}(k_i \mid i \leftrightarrow j)$ be the **conditional probability** for node i to have degree k_i given it is connected to some random other node j.
- Suppose we have a **random network** with a **GC**. Assuming that a node connected to at least one other node is actually linked to the GC we obtain

$$\langle k \rangle_{\mathrm{G}} \simeq \langle k_i \mid i \leftrightarrow j \rangle = \sum_i k_i \mathcal{P}(k_i \mid i \leftrightarrow j) \geq 2.$$

Condition for the existence of a giant component

- Assume a network with a giant component.
- The average degree in a giant component: $\langle k \rangle_G \geq 2$

$\leftrightarrow$

- Let $\mathcal{P}(k_i \mid i \leftrightarrow j)$ be the **conditional probability** for node i to have degree k_i given it is connected to some random other node j.

- Suppose we have a **random network** with a **GC**. Assuming that a node connected to at least one other node is actually linked to the GC we obtain

$$\langle k \rangle_G \simeq \langle k_i \mid i \leftrightarrow j \rangle = \sum_i k_i \mathcal{P}(k_i \mid i \leftrightarrow j) \geq 2.$$

Condition for the existence of a giant component

- Using the definition of the conditional probability,

$$\mathcal{P}(k_i \mid i \leftrightarrow j) = \frac{\mathcal{P}(k_i, i \leftrightarrow j)}{\mathcal{P}(i \leftrightarrow j)}$$

- In general, for random networks

$$\mathcal{P}(i \leftrightarrow j \mid k_i) = \frac{k_i}{N}, \qquad \mathcal{P}(i \leftrightarrow j) = \frac{2M}{N(N-1)} = \frac{\langle k \rangle}{N}$$

→ Thus, the condition for the existence of a GC:

$$\sum_i k_i \mathcal{P}(k_i \mid i \leftrightarrow j) = \sum_i k_i \frac{\mathcal{P}(i \leftrightarrow j \mid k_i) p(k_i)}{\mathcal{P}(i \leftrightarrow j)} =$$

$$\sum_i k_i \frac{k_i p(k_i)}{\langle k \rangle} = \frac{\sum_i k_i^2 p(k_i)}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle} \geq 2.$$

Condition for the existence of a giant component

- Using the definition of the conditional probability,

$$\mathcal{P}(k_i \mid i \leftrightarrow j) = \frac{\mathcal{P}(k_i, i \leftrightarrow j)}{\mathcal{P}(i \leftrightarrow j)}$$

$$\rightarrow \quad \mathcal{P}(k_i \mid i \leftrightarrow j)\mathcal{P}(i \leftrightarrow j) = \mathcal{P}(k_i, i \leftrightarrow j)$$

- In general, for random networks

$$\mathcal{P}(i \leftrightarrow j \mid k_i) = \frac{k_i}{N}, \qquad \mathcal{P}(i \leftrightarrow j) = \frac{2M}{N(N-1)} = \frac{\langle k \rangle}{N}$$

→ Thus, the condition for the existence of a GC:

$$\sum_i k_i \mathcal{P}(k_i \mid i \leftrightarrow j) = \sum_i k_i \frac{\mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)}{\mathcal{P}(i \leftrightarrow j)} =$$

$$\sum_i k_i \frac{k_i p(k_i)}{\langle k \rangle} = \frac{\sum_i k_i^2 p(k_i)}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle} \geq 2.$$

Condition for the existence of a giant component

- Using the definition of the conditional probability,

$$\mathcal{P}(k_i \mid i \leftrightarrow j) = \frac{\mathcal{P}(k_i, i \leftrightarrow j)}{\mathcal{P}(i \leftrightarrow j)}$$

$$\rightarrow \quad \mathcal{P}(k_i \mid i \leftrightarrow j)\mathcal{P}(i \leftrightarrow j) = \mathcal{P}(k_i, i \leftrightarrow j) = \mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)$$

- In general, for random networks

$$\mathcal{P}(i \leftrightarrow j \mid k_i) = \frac{k_i}{N}, \qquad \mathcal{P}(i \leftrightarrow j) = \frac{2M}{N(N-1)} = \frac{\langle k \rangle}{N}$$

→ Thus, the condition for the existence of a GC:

$$\sum_i k_i \mathcal{P}(k_i \mid i \leftrightarrow j) = \sum_i k_i \frac{\mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)}{\mathcal{P}(i \leftrightarrow j)} =$$

$$\sum_i k_i \frac{k_i p(k_i)}{\langle k \rangle} = \frac{\sum_i k_i^2 p(k_i)}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle} \geq 2.$$

Condition for the existence of a giant component

- Using the definition of the conditional probability,

$$\mathcal{P}(k_i \mid i \leftrightarrow j) = \frac{\mathcal{P}(k_i, i \leftrightarrow j)}{\mathcal{P}(i \leftrightarrow j)}$$

$$\rightarrow \quad \mathcal{P}(k_i \mid i \leftrightarrow j)\mathcal{P}(i \leftrightarrow j) = \mathcal{P}(k_i, i \leftrightarrow j) = \mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)$$

$$\rightarrow \quad \mathcal{P}(k_i \mid i \leftrightarrow j) = \frac{\mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)}{\mathcal{P}(i \leftrightarrow j)}$$

- In general, for random networks

$$\mathcal{P}(i \leftrightarrow j \mid k_i) = \frac{k_i}{N}, \qquad \mathcal{P}(i \leftrightarrow j) = \frac{2M}{N(N-1)} = \frac{\langle k \rangle}{N}$$

→ Thus, the condition for the existence of a GC:

$$\sum_i k_i \mathcal{P}(k_i \mid i \leftrightarrow j) = \sum_i k_i \frac{\mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)}{\mathcal{P}(i \leftrightarrow j)} =$$

$$\sum_i k_i \frac{k_i p(k_i)}{\langle k \rangle} = \frac{\sum_i k_i^2 p(k_i)}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle} \geq 2.$$

Condition for the existence of a giant component

- Using the definition of the conditional probability,

$$\mathcal{P}(k_i \mid i \leftrightarrow j) = \frac{\mathcal{P}(k_i, i \leftrightarrow j)}{\mathcal{P}(i \leftrightarrow j)}$$

$$\rightarrow \quad \mathcal{P}(k_i \mid i \leftrightarrow j)\mathcal{P}(i \leftrightarrow j) = \mathcal{P}(k_i, i \leftrightarrow j) = \mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)$$

$$\rightarrow \quad \mathcal{P}(k_i \mid i \leftrightarrow j) = \frac{\mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)}{\mathcal{P}(i \leftrightarrow j)}$$

- In general, for random networks

$$\mathcal{P}(i \leftrightarrow j \mid k_i) = \frac{k_i}{N}, \qquad \mathcal{P}(i \leftrightarrow j) = \frac{2M}{N(N-1)} = \frac{\langle k \rangle}{N}$$

→ Thus, the condition for the existence of a GC:

$$\sum_i k_i \mathcal{P}(k_i \mid i \leftrightarrow j) = \sum_i k_i \frac{\mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)}{\mathcal{P}(i \leftrightarrow j)} =$$

$$\sum_i k_i \frac{k_i p(k_i)}{\langle k \rangle} = \frac{\sum_i k_i^2 p(k_i)}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle} \geq 2.$$

Condition for the existence of a giant component

- Using the definition of the conditional probability,

$$\mathcal{P}(k_i \mid i \leftrightarrow j) = \frac{\mathcal{P}(k_i, i \leftrightarrow j)}{\mathcal{P}(i \leftrightarrow j)}$$

$$\rightarrow \quad \mathcal{P}(k_i \mid i \leftrightarrow j)\mathcal{P}(i \leftrightarrow j) = \mathcal{P}(k_i, i \leftrightarrow j) = \mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)$$

$$\rightarrow \quad \mathcal{P}(k_i \mid i \leftrightarrow j) = \frac{\mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)}{\mathcal{P}(i \leftrightarrow j)}$$

- In general, for random networks

$$\mathcal{P}(i \leftrightarrow j \mid k_i) = \frac{k_i}{N}, \qquad \mathcal{P}(i \leftrightarrow j) = \frac{2M}{N(N-1)} = \frac{\langle k \rangle}{N}$$

→ Thus, the condition for the existence of a GC:

$$\sum_i k_i \mathcal{P}(k_i \mid i \leftrightarrow j) = \sum_i k_i \frac{\mathcal{P}(i \leftrightarrow j \mid k_i)p(k_i)}{\mathcal{P}(i \leftrightarrow j)} =$$

$$\sum_i k_i \frac{k_i p(k_i)}{\langle k \rangle} = \frac{\sum_i k_i^2 p(k_i)}{\langle k \rangle} = \frac{\langle k^2 \rangle}{\langle k \rangle} \geq 2.$$

- Accordingly, an E-R graph has a GC if

$$p(k) \simeq \frac{\langle k\rangle^k e^{-\langle k\rangle}}{k!}, \rightarrow \left\langle k^2\right\rangle = \langle k\rangle\left(1+\langle k\rangle\right),$$

$$\rightarrow \frac{\left\langle k^2\right\rangle}{\langle k\rangle} = 1+\langle k\rangle \geq 2$$

$$\rightarrow \langle k\rangle \geq 1.$$

(The same result was derived in different way in the 2nd part of the lecture.)

Condition for the existence of a giant component

What about **scale-free** networks?

Condition for the existence of a giant component

What about **scale-free** networks?

- For most real networks $p(k) \sim k^{-\gamma}$ with $2 \leq \gamma \leq 3$,

$\rightarrow$ $\langle k^2 \rangle$ is diverging!

Condition for the existence of a giant component

What about **scale-free** networks?

- For most real networks $p(k) \sim k^{-\gamma}$ with $2 \leq \gamma \leq 3$,

$\rightarrow$ $\langle k^2 \rangle$ is diverging!

$\rightarrow$ Thus, according to the $\frac{\langle k^2 \rangle}{\langle k \rangle} \geq 2$ condition, scale-free networks have a GC independent of the average degree.

Critical f

- Now that we have this simple condition for the existence of the giant component, we can examine the random node removal process on an analytical ground as follows:

Critical f

- Now that we have this simple condition for the existence of the giant component, we can examine the random node removal process on an analytical ground as follows:
  - We assume the removal of f fraction of the nodes at random.

Critical f

- Now that we have this simple condition for the existence of the giant component, we can examine the random node removal process on an analytical ground as follows:
  - We assume the removal of f fraction of the nodes at random.
  - Based on the original $p(k)$ of the network and f, we calculate the modified $p'(k')$.

Critical f

- Now that we have this simple condition for the existence of the giant component, we can examine the random node removal process on an analytical ground as follows:
  - We assume the removal of f fraction of the nodes at random.
  - Based on the original $p(k)$ of the network and f, we calculate the modified $p'(k')$.
  - We check whether the condition is fulfilled or not according to $p'(k')$ or not.

Critical f

- Now that we have this simple condition for the existence of the giant component, we can examine the random node removal process on an analytical ground as follows:
  - We assume the removal of f fraction of the nodes at random.
  - Based on the original $p(k)$ of the network and f, we calculate the modified $p'(k')$.
  - We check whether the condition is fulfilled or not according to $p'(k')$ or not.
  - The smallest possible value of f for which the condition is not fulfilled defines the critical f, where we reach the transition point.

Network breakdown

Assuming the random removal of f fraction of nodes from a network characterized by a given $p(k)$ (and corresponding $\langle k \rangle$, $\langle k^2 \rangle$):

Network breakdown

Assuming the random removal of f fraction of nodes from a network characterized by a given $p(k)$ (and corresponding $\langle k \rangle, \langle k^2 \rangle$):

We obtain a modified $p'(k')$ with modified $\langle k' \rangle, \langle (k')^2 \rangle$ and have to examine whether

$$\frac{\langle (k')^2 \rangle}{\langle k' \rangle} \overset{?}{\geq} 2.$$

Network breakdown

Assuming the random removal of f fraction of nodes from a network characterized by a given $p(k)$ (and corresponding $\langle k\rangle, \langle k^2\rangle$):

We obtain a modified $p'(k')$ with modified $\langle k'\rangle, \langle (k')^2\rangle$ and have to examine whether

$$\frac{\langle (k')^2\rangle}{\langle k'\rangle} \overset{?}{\geq} 2.$$

- Simple estimate for $\langle k'\rangle$ after the removal if $f \ll 1$?

Network breakdown

Assuming the random removal of f fraction of nodes from a network characterized by a given $p(k)$ (and corresponding $\langle k\rangle, \langle k^2\rangle$):

We obtain a modified $p'(k')$ with modified $\langle k'\rangle, \langle (k')^2\rangle$ and have to examine whether

$$\frac{\langle (k')^2\rangle}{\langle k'\rangle} \overset{?}{\geq} 2.$$

- Simple estimate for $\langle k'\rangle$ after the removal if $f \ll 1$:
  - f fraction of removed nodes induces f fraction of removed links,

  $\rightarrow$ in other words, $1-f$ fraction of the links remain, giving $\langle k'\rangle = (1-f)\langle k\rangle$.

Network breakdown

Assuming the random removal of f fraction of nodes from a network characterized by a given $p(k)$ (and corresponding $\langle k\rangle, \langle k^2\rangle$):

Network breakdown

Assuming the random removal of f fraction of nodes from a network characterized by a given $p(k)$ (and corresponding $\langle k\rangle, \langle k^2\rangle$):

- A node with originally k degree will lose q neighbors with probability

$$\mathcal{P}_k(\text{loose } q \text{ links}) = \binom{k}{q} f^q (1-f)^{k-q}$$

Network breakdown

Assuming the random removal of f fraction of nodes from a network characterized by a given $p(k)$ (and corresponding $\langle k\rangle, \langle k^2\rangle$):

- A node with originally k degree will lose q neighbors with probability

$$\mathcal{P}_k(\text{loose } q \text{ links}) = \binom{k}{q} f^q (1-f)^{k-q}$$

- Its new degree shall be $k' = k - q$, thus,

$$\mathcal{P}(k \to k') = \binom{k}{q} f^q (1-f)^{k-q} = \binom{k}{k-k'} f^{k-k'} (1-f)^{k'} = \binom{k}{k'} f^{k-k'} (1-f)^{k'}.$$

Network breakdown

Assuming the random removal of f fraction of nodes from a network characterized by a given $p(k)$ (and corresponding $\langle k\rangle, \langle k^2\rangle$):

- A node with originally k degree will lose q neighbors with probability

$$\mathcal{P}_k(\text{loose } q \text{ links}) = \binom{k}{q} f^q (1-f)^{k-q}$$

- Its new degree shall be $k' = k - q$, thus,

$$\mathcal{P}(k \to k') = \binom{k}{q} f^q (1-f)^{k-q} = \binom{k}{k-k'} f^{k-k'} (1-f)^{k'} = \binom{k}{k'} f^{k-k'} (1-f)^{k'}.$$

- Therefore, the degree distribution after the removal is

$$p'(k') = \sum_{k=k'}^{\infty} p(k) \binom{k}{k'} f^{k-k'} (1-f)^{k'}.$$

The critical f

- The modified $\langle k' \rangle$ after the removal:

$$\begin{aligned}\langle k' \rangle &= \sum_{k'=0}^{\infty} k' p'(k') = \sum_{k'=0}^{\infty} k' \sum_{k=k'}^{\infty} p(k) \binom{k}{k'} f^{k-k'} (1-f)^{k'} = \\ & \sum_{k'=0}^{\infty} \sum_{k=k'}^{\infty} p(k) \frac{k(k-1)!}{(k'-1)!(k-k')!} f^{k-k'} (1-f)^{k'-1} (1-f)\end{aligned}$$

The sum is done over the triangle, thus, it can be replaced as $\sum_{k'=0}^{\infty} \sum_{k=k'}^{\infty} = \sum_{k=0}^{\infty} \sum_{k'=0}^{k}$

$$\begin{aligned}\langle k' \rangle &= \sum_{k=0}^{\infty} \sum_{k'=0}^{k} p(k) \frac{k(k-1)!}{(k'-1)!(k-k')!} f^{k-k'} (1-f)^{k'-1} (1-f) = \\ & (1-f) \sum_{k=0}^{\infty} k p(k) \underbrace{\sum_{k'=0}^{k} \binom{k-1}{k'-1} f^{k-k'} (1-f)^{k'-1}}_{1} = (1-f) \langle k \rangle\end{aligned}$$

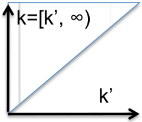
The critical f

- The modified $\langle k' \rangle$ after the removal:

$$\langle k' \rangle = \sum_{k'=0}^{\infty} k' p'(k') = \sum_{k'=0}^{\infty} k' \sum_{k=k'}^{\infty} p(k) \binom{k}{k'} f^{k-k'} (1-f)^{k'} =$$

$$\sum_{k'=0}^{\infty} \sum_{k=k'}^{\infty} p(k) \frac{k(k-1)!}{(k'-1)!(k-k')!} f^{k-k'} (1-f)^{k'-1} (1-f)$$

The sum is done over the triangle, thus, it can be replaced as $\sum_{k'=0}^{\infty} \sum_{k=k'}^{\infty} = \sum_{k=0}^{\infty} \sum_{k'=0}^{k}$



$$\langle k' \rangle = \sum_{k=0}^{\infty} \sum_{k'=0}^{k} p(k) \frac{k(k-1)!}{(k'-1)!(k-k')!} f^{k-k'} (1-f)^{k'-1} (1-f) =$$

$$(1-f) \sum_{k=0}^{\infty} k p(k) \underbrace{\sum_{k'=0}^{k} \binom{k-1}{k'-1} f^{k-k'} (1-f)^{k'-1}}_{1} = (1-f) \langle k \rangle$$

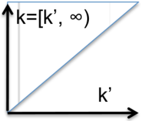
The critical f

- The modified $\langle k' \rangle$ after the removal:

$$\langle k' \rangle = \sum_{k'=0}^{\infty} k' p'(k') = \sum_{k'=0}^{\infty} k' \sum_{k=k'}^{\infty} p(k) \binom{k}{k'} f^{k-k'} (1-f)^{k'} =$$

$$\sum_{k'=0}^{\infty} \sum_{k=k'}^{\infty} p(k) \frac{k(k-1)!}{(k'-1)!(k-k')!} f^{k-k'} (1-f)^{k'-1} (1-f)$$

The sum is done over the triangle, thus, it can be replaced as $\sum_{k'=0}^{\infty} \sum_{k=k'}^{\infty} = \sum_{k=0}^{\infty} \sum_{k'=0}^{k}$



$$\langle k' \rangle = \sum_{k=0}^{\infty} \sum_{k'=0}^{k} p(k) \frac{k(k-1)!}{(k'-1)!(k-k')!} f^{k-k'} (1-f)^{k'-1} (1-f) =$$

$$(1-f) \sum_{k=0}^{\infty} k p(k) \underbrace{\sum_{k'=0}^{k} \binom{k-1}{k'-1} f^{k-k'} (1-f)^{k'-1}}_{1} = (1-f) \langle k \rangle$$

The critical f

- The modified $\langle (k')^2 \rangle$:

$$\langle (k')^2 \rangle = \langle k'(k'-1) - k' \rangle = \langle k'(k'-1) \rangle - \langle k' \rangle \, .$$

$$\langle k'(k'-1) \rangle =$$

$$\sum_{k'=0}^{\infty} \sum_{k=k'}^{\infty} p(k) \frac{k(k-1)(k-2)!}{(k'-2)!(k-k')!} f^{k-k'} (1-f)^{k'-2} (1-f)^2 =$$

$$(1-f)^2 \sum_{k=0}^{\infty} \sum_{k'=0}^{k} k(k-1)p(k) \frac{(k-2)!}{(k'-2)!(k-k')!} f^{k-k'} (1-f)^{k'-2} =$$

$$(1-f)^2 \sum_{k=0}^{\infty} k(k-1)p(k) \underbrace{\sum_{k'=0}^{k} \binom{k-2}{k'-2} f^{k-k'} (1-f)^{k'-2}}_{1} =$$

$$= (1-f)^2 \langle k(k-1) \rangle \, .$$

The critical f

- The modified $\left\langle (k')^2 \right\rangle$:

$$\left\langle (k')^2 \right\rangle = \left\langle k'(k'-1) - k' \right\rangle = \left\langle k'(k'-1) \right\rangle - \left\langle k' \right\rangle .$$

$$\left\langle k'(k'-1) \right\rangle =$$

$$\sum_{k'=0}^{\infty} \sum_{k=k'}^{\infty} p(k) \frac{k(k-1)(k-2)!}{(k'-2)!(k-k')!} f^{k-k'} (1-f)^{k'-2} (1-f)^2 =$$

$$(1-f)^2 \sum_{k=0}^{\infty} \sum_{k'=0}^{k} k(k-1)p(k) \frac{(k-2)!}{(k'-2)!(k-k')!} f^{k-k'} (1-f)^{k'-2} =$$

$$(1-f)^2 \sum_{k=0}^{\infty} k(k-1)p(k) \underbrace{\sum_{k'=0}^{k} \binom{k-2}{k'-2} f^{k-k'} (1-f)^{k'-2}}_{1} =$$

$$= (1-f)^2 \left\langle k(k-1) \right\rangle .$$

The critical f

- The modified $\left\langle (k')^2 \right\rangle$:

$$\begin{aligned}\left\langle (k')^2 \right\rangle &= (1-f)^2 \langle k(k-1)\rangle - \left\langle k' \right\rangle = \\ &(1-f)^2 \left(\left\langle k^2 \right\rangle - \langle k \rangle\right) - (1-f)\langle k \rangle = \\ &(1-f)^2 \left\langle k^2 \right\rangle + f(1-f)\langle k \rangle .\end{aligned}$$

- The condition for having a GC:

$$\begin{aligned}\frac{\left\langle (k')^2 \right\rangle}{\left\langle k' \right\rangle} &= \frac{(1-f)^2 \left\langle k^2 \right\rangle + f(1-f)\langle k \rangle}{(1-f)\langle k \rangle} = (1-f)\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} + f, \\ &\to \frac{\left\langle k^2 \right\rangle}{\langle k \rangle} + f\left(1 - \frac{\left\langle k^2 \right\rangle}{\langle k \rangle}\right) \geq 2\end{aligned}$$

- The critical fraction of removed links:

$$f_c = \frac{2 - \frac{\left\langle k^2 \right\rangle}{\langle k \rangle}}{1 - \frac{\left\langle k^2 \right\rangle}{\langle k \rangle}} = \frac{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 2}{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 1} = 1 - \frac{1}{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 1}$$

The critical f

- The modified $\left\langle (k')^2 \right\rangle$:

$$\begin{aligned}\left\langle (k')^2 \right\rangle &= (1-f)^2 \langle k(k-1)\rangle - \langle k' \rangle = \\ &(1-f)^2 \left(\left\langle k^2 \right\rangle - \langle k \rangle\right) - (1-f)\langle k \rangle = \\ &(1-f)^2 \left\langle k^2 \right\rangle + f(1-f)\langle k \rangle .\end{aligned}$$

- The condition for having a GC:

$$\begin{aligned}\frac{\left\langle (k')^2 \right\rangle}{\langle k' \rangle} &= \frac{(1-f)^2 \left\langle k^2 \right\rangle + f(1-f)\langle k \rangle}{(1-f)\langle k \rangle} = (1-f)\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} + f, \\ &\rightarrow \frac{\left\langle k^2 \right\rangle}{\langle k \rangle} + f\left(1 - \frac{\left\langle k^2 \right\rangle}{\langle k \rangle}\right) \geq 2\end{aligned}$$

- The critical fraction of removed links:

$$f_c = \frac{2 - \frac{\left\langle k^2 \right\rangle}{\langle k \rangle}}{1 - \frac{\left\langle k^2 \right\rangle}{\langle k \rangle}} = \frac{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 2}{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 1} = 1 - \frac{1}{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 1}$$

The critical f

- The modified $\left\langle (k')^2 \right\rangle$:

$$\begin{aligned}\left\langle (k')^2 \right\rangle &= (1-f)^2 \langle k(k-1) \rangle - \left\langle k' \right\rangle = \\ &(1-f)^2 \left(\left\langle k^2 \right\rangle - \langle k \rangle \right) - (1-f) \langle k \rangle = \\ &(1-f)^2 \left\langle k^2 \right\rangle + f(1-f) \langle k \rangle .\end{aligned}$$

- The condition for having a GC:

$$\begin{aligned}\frac{\left\langle (k')^2 \right\rangle}{\left\langle k' \right\rangle} &= \frac{(1-f)^2 \left\langle k^2 \right\rangle + f(1-f) \langle k \rangle}{(1-f) \langle k \rangle} = (1-f) \frac{\left\langle k^2 \right\rangle}{\langle k \rangle} + f, \\ &\to \frac{\left\langle k^2 \right\rangle}{\langle k \rangle} + f \left(1 - \frac{\left\langle k^2 \right\rangle}{\langle k \rangle} \right) \geq 2\end{aligned}$$

- The critical fraction of removed links:

$$f_c = \frac{2 - \frac{\left\langle k^2 \right\rangle}{\langle k \rangle}}{1 - \frac{\left\langle k^2 \right\rangle}{\langle k \rangle}} = \frac{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 2}{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 1} = 1 - \frac{1}{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 1}$$

The critical f

The critical fraction of removed links: $f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$

- For an **E-R graph**:

$$\begin{aligned} \langle k^2 \rangle &= \langle k \rangle (1 + \langle k \rangle), \\ \rightarrow f_c &= 1 - \frac{1}{\langle k \rangle}, \\ \rightarrow \langle k' \rangle_{f_c} &= (1 - f_c) \langle k \rangle = 1. \end{aligned}$$

Thus, in order to destroy an E-R graph we have to keep on removing the links until reaching $\langle k' \rangle = 1$.

The critical f

The critical fraction of removed links: $f_c = 1 - \frac{1}{\frac{\left\langle k^2 \right\rangle}{\langle k \rangle} - 1}$

- For an **E-R graph**:

$$
\begin{aligned}
\left\langle k^2 \right\rangle &= \langle k \rangle \left(1 + \langle k \rangle\right), \\
\to f_c &= 1 - \frac{1}{\langle k \rangle}, \\
\to \left\langle k' \right\rangle_{f_c} &= \left(1 - f_c\right) \langle k \rangle = 1.
\end{aligned}
$$

Thus, in order to destroy an E-R graph we have to keep on removing the links until reaching $\langle k' \rangle = 1$.

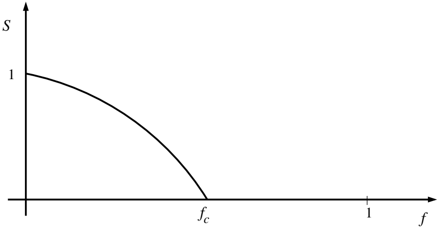
The critical f

The critical fraction of removed links: $f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$

- For an **E-R graph**:

$$\begin{aligned} \langle k^2 \rangle &= \langle k \rangle (1 + \langle k \rangle), \\ \rightarrow f_c &= 1 - \frac{1}{\langle k \rangle}, \\ \rightarrow \langle k' \rangle_{f_c} &= (1 - f_c) \langle k \rangle = 1. \end{aligned}$$

Thus, in order to destroy an E-R graph we have to keep on removing the links until reaching $\langle k' \rangle = 1$.



The critical f

The critical fraction of removed links: $f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$

- For a **scale-free graph**:

$$\begin{aligned} \langle k^2 \rangle &\to \infty \quad (\text{if } 2 \leq \gamma \leq 3) \\ \to f_c &\to 1 \end{aligned}$$

Thus, **scale-free networks are extremely robust against random node (link) removal!**

The critical f

The critical fraction of removed links: $f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$

- For a **scale-free graph**:

$$\begin{aligned} \langle k^2 \rangle &\rightarrow \infty \quad (\text{if } 2 \leq \gamma \leq 3) \\ \rightarrow f_c &\rightarrow 1 \end{aligned}$$

Thus, **scale-free networks are extremely robust against random node (link) removal!**

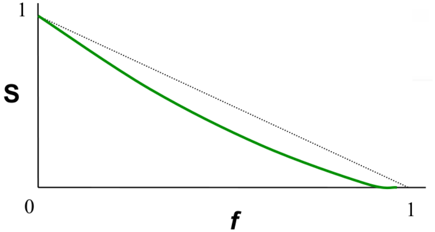
The critical f

The critical fraction of removed links: $f_c = 1 - \dfrac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$

- For a **scale-free graph**:

$$\begin{aligned} \langle k^2 \rangle &\to \infty \text{ (if } 2 \leq \gamma \leq 3) \\ \to f_c &\to 1 \end{aligned}$$

Thus, **scale-free networks are extremely robust against random node (link) removal!**

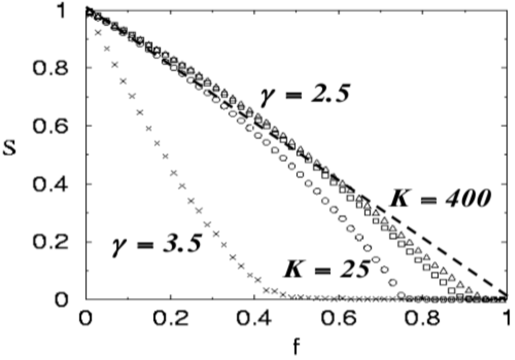


Extreme robustness of scale-free networks

Real world examples:

- Internet, router level map: $N = 228,263$, $\gamma = 2.1 \pm 0.1$
  $\rightarrow f_c = 0.962$
- Internet, AS level map: $N = 11,164$, $\gamma = 2.1 \pm 0.1$
  $\rightarrow f_c = 0.996$

Figure from the paper by Barabási and co-workers:



Extreme robustness of scale-free networks

Why do we see scale-free networks everywhere?

→ A very plausible answer is their extreme robustness!

The price of robustness

This extreme robustness has a price...

The price of robustness

This extreme robustness has a price...

Scale-free networks are vulnerable to targeted attack!

Attack tolerance

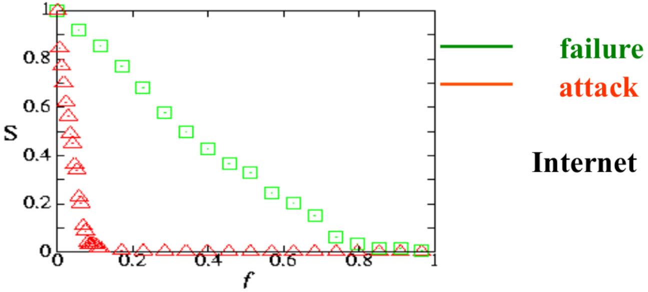
How to “attack” a scale-free network?

→ Removing nodes in the order of their degree.

Attack tolerance

How to “attack” a scale-free network?

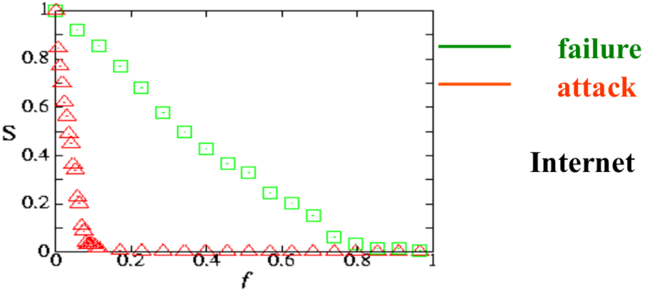
→ Removing nodes in the order of their degree.



Attack tolerance

How to “attack” a scale-free network?

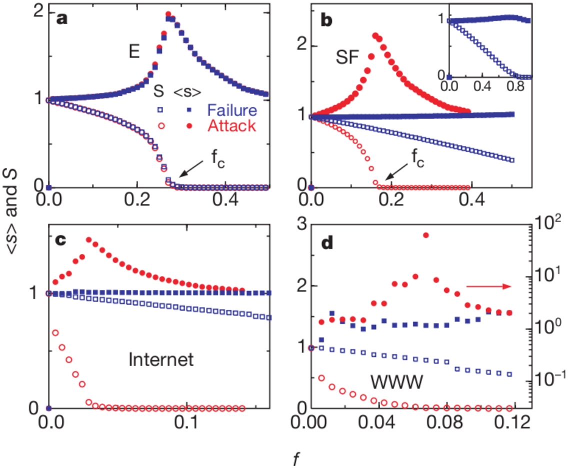
→ Removing nodes in the order of their degree.



Please take a look at the accompanying jupyter notebook where we simulate also attack tolerance on networks.

Error and attack tolerance

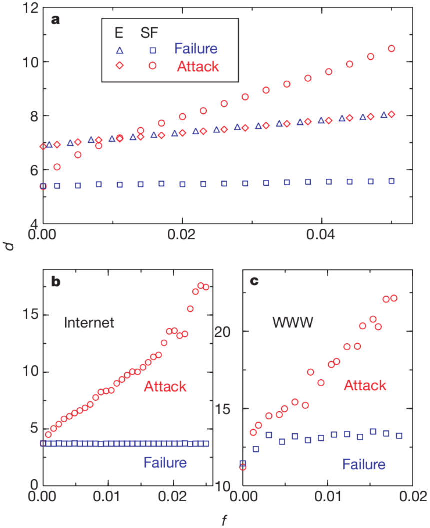
Figure from the paper by Albert et. al., showing the behaviour of S and χ:



R. Albert, H. Jeong and A.-L. Barabási, *Nature* **406**, 378 (2000)

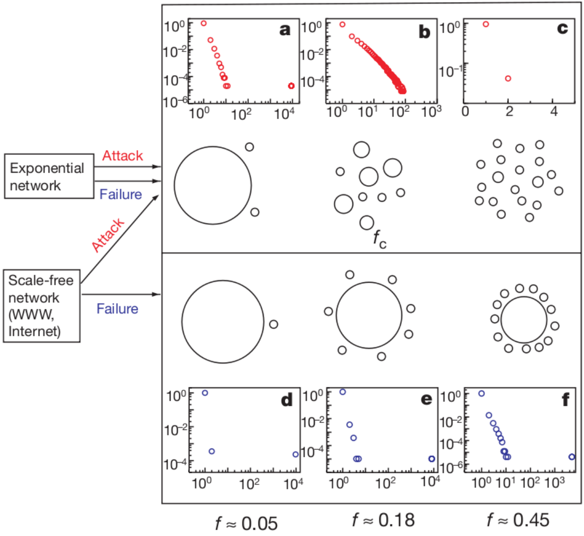
Error and attack tolerance

Figure from the paper by Albert et. al., showing the behaviour of the average distance:



Error and attack tolerance

Summary figure from the paper by Albert et. al.:



Summary

Network robustness short summary

- Scale-free networks are extremely robust against random node removal. This property is connected to the divergence of the second moment of the degree distribution.
- However, in the mean time scale-free networks are also fragile against targeted attack, since taking out the HUBS is producing a lot of damage.
- We examined these properties with the help of the giant component, within the framework of percolation transitions.

Further reading on network robustness (not compulsory):

Network science book by A.-L. Barabási, chapter 8:
`http://networksciencebook.com/chapter/8`