

# Network Science (VU) (706.703)

## Function of Networks

Denis Helic

ISDS, TU Graz

January 24, 2018

# Outline

- 1 Introduction
- 2 Percolation
- 3 Spread Processes
- 4 Dynamical Systems

# Structure and function of networks

- One of the main goals of studying networks is to understand the behavior of the systems we represent with the networks
- E.g. we study online social network to understand how information and ideas spread through these networks
- We study technological networks to analyze the protocols and improve them
- We study metabolic networks to understand chemical process in the cell
- We study information networks to understand the user behavior

# Structure and function of networks

- Studying the structure is just the first step towards understanding
- Another important step is to make the connection between the structure and the function
- Once when we have measured and quantified the structure the question is then how can we turn these results into prediction of the system's behavior?
- There are some areas where some results have been achieved
- Resilience, spreading processes, and dynamical systems on networks

# Structure and function of networks

- Resilience
  - Attack the networks and see how resilient are they: percolation
- Spreading processes
  - Disease spreading, computer viruses, ideas, innovation: threshold and dynamics
- Dynamical systems
  - Diffusion, opinion dynamics, activity: dynamics

# Percolation

- Imagine removing some fraction of nodes, along with its links from a network
- This process is called percolation (site percolation)
- This simple process can be used as a model of node failures
- E.g. router failures on the Internet
- E.g. vaccination of individuals against the disease

# Percolation

- Sometimes links in a network fail
- E.g. communication lines on the Internet fail
- We can model this by removing some fraction of links from the network (bond percolation)
- In general, there are many ways how nodes and links can be removed from the network
- The simplest model is to remove them uniformly at random
- Another possibility is to remove them according to their degrees

# Uniform random removal of nodes

- The percolation process is parametrized by probability  $\phi$ , which is the probability that a node is present in the network
- If present the node is said to be *occupied*
- $\phi$  is called occupation probability
- Thus, if  $\phi = 1$  none of the nodes has been removed
- If  $\phi = 0$  all nodes have been removed



# Uniform random removal of nodes

- When  $\phi$  is close to one, the nodes tend to be connected and form a giant component
- When  $\phi$  is close to zero, there are only few nodes that are typically disconnected
- When we start with a large  $\phi$  and decrease it there comes a point where the giant component breaks up
- When we start with a small  $\phi$  and increase it, there comes a point where the small and disconnected components are connected to form a giant component
- This point is called percolation threshold

# Non-uniform random removal of nodes

- Typically, in this model  $\phi$  depends on the degree of a node
- E.g. it might be proportional to the degree and this is biased towards removal of the hubs
- Or it can be inversely proportional to the degree, i.e. we tend to remove low degree nodes first
- The basic process and formation of the giant component still take place
- But, the percolation threshold might change

# Percolation in real-world networks

- Typically, we study percolation by simulating node removal
- We are interested in the size of the largest component as the function of  $\phi$
- We should repeat simulation many times, i.e. thousand times to see what is the typical behavior
- We are also interested in variation around that typical behavior

# Percolation in real-world networks

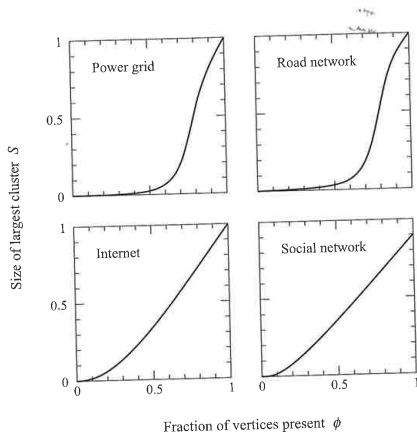


Figure: Uniform removal

# Percolation in real-world networks

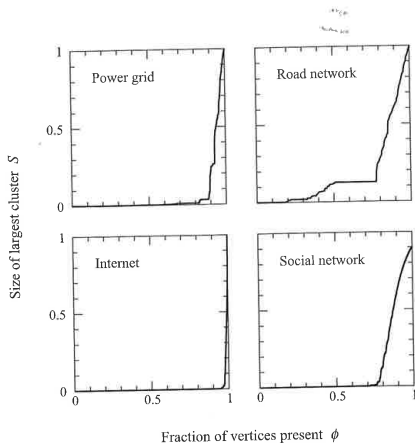


Figure: Non-uniform (targeted) removal

# Spread over networks

- Connection of social networks with the spread of disease
- Diseases spread over networks of contacts between individuals
- Similarly, the computer viruses spread over e.g. the Web, Facebook, e-mail communication, etc.
- The mechanics are similar
- Also, information diffusion, gossip, spread of ideas have similar underlying (abstract) dynamics
- Similar models can be used to describe all of these processes

# Spread over networks

- The biological process that leads to infection is complicated
- E.g. multiplication of viruses, bacteria
- The interplay with the immune system
- Previous diseases
- The outcome of the diseases, etc.
- If we want to fully understand this process we need to understand the biology behind it

# Spread over networks

- Similarly, the process of how ideas or even information spreads and gets adopted is complicated
- Intrinsic personal tendency to accept new ideas
- Based on education, previous knowledge, previous experience, etc.
- Attractiveness of the idea itself
- The interplay with the social context, etc.
- If we want to fully understand this process we need to understand the psychology, sociology, group psychology, etc.



# Models of the spread

- A standard scientific approach is to start with a simplified model of reality
- There are many different approaches to modeling
- A possible categorization:
  - 1 Threshold models
  - 2 Agent-based models
  - 3 Dynamics

# Models of the spread

- Another categorization:
  - 1 Deterministic models
  - 2 Probabilistic models
- Also we can distinguish between:
  - 1 Discrete models
  - 2 Continuous models

# Threshold model

- E.g. a threshold model such as linear threshold model by Granovetter
- Each person has an individual threshold (some number)
- If a certain value goes over threshold the person engages in a behavior
- E.g. adopts an idea, or information

# Threshold model

- The threshold depends on social context, education, age, personality, etc.
- It abstracts all of these factors with a single number
- Threshold model is a standard deterministic model
- I.e. whenever we pass the threshold the behavior is triggered

# Threshold model

- Alternatively, we can also have a probabilistic threshold model
- I.e. a probabilistic model
- We define a probability that a person engages in behavior and draw random numbers
- The probability is based on the same variables as threshold
- However, the emerging behavior is probabilistic

# Agent-based models

- Each node is an agent that acts autonomously
- I.e. each agent has a certain behavior
- Agents engage in interactions with other agents
- E.g. neighbors with neighbors or sometimes with arbitrary other agents
- The system evolution is defined by the rules of the interactions

# Naming game

- How ideas are spread and adopted in a network
- How does consensus emerge?
- We have a number of people in a social network that are discussing some problems and are trying to reach a consensus
- We model people and their relationships as nodes and links in a social network
- Each person carries a list of words, e.g. names, solutions for the issue that is being discussed

# Naming game

- The rules of the game:
  - ① We select a random person (speaker) from the network
  - ② Then we select a random neighbor (listener) of that person
  - ③ The the speaker selects a random word from her name list
  - ④ If the listener has this word in her list they both agree on this word, keep it and throw all other words from their lists
  - ⑤ If the listener does not possess the word then she includes it in her list
  - ⑥ We repeat the game many times
- The Naming Game in Social Networks: Community Formation and Consensus Engineering: <http://arxiv.org/abs/1008.1096>



# Naming game

- Demo
- Some results:
  - 1 The network structure is crucial for the process dynamics
  - 2 E.g. community structure
  - 3 It is possible to manipulate (engineer) the dynamics by e.g. committed agents
  - 4 Danger: manipulation of public opinion, mass media, etc.

# Dynamical systems

- A general concept of dynamical systems
- We represent the state of a dynamical system by some set of quantitative variables
- Each node has a variable  $x_i(t)$
- E.g. the probability of being infected, the number of people waiting on the bus station

# Dynamical systems

- A *dynamical system* is any system whose *state changes over time*
- According to some given rules or equations
- E.g. equations of change of probabilities of being infected in the SI model
- E.g. equations of how people arrive and leave bus stations

# Dynamical systems

- Many other processes (or simplified versions of these processes) can be represented as dynamical systems
- The spread of news or information between friends
- The movement of money through economy
- The flow of traffic on roads
- Data over the Internet, etc.

# Dynamical systems

- For the rest of our discussion on the function of complex networks we will concentrate on such models
- But for this we yet again need some mathematical basics and background
- Differential equations