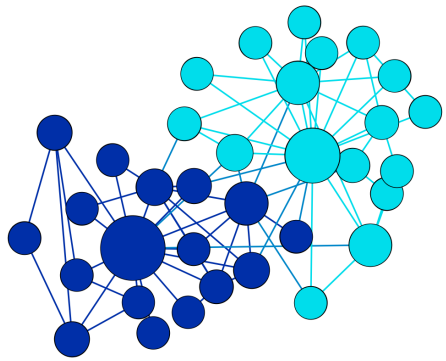
Analysis of Social and Information Networks (PRODEI040)

Pedro Ribeiro

DCC/FCUP

2020/2021



Who am I?

Name: Pedro Ribeiro
Office 1.47 (FC6 - DCC building)
Personal Website: http://www.dcc.fc.up.pt/~pribeiro/

PhD in **Computer Science**

Main research interests:

- Complex Network Analysis, Network Science, Graph Mining, Data Mining.
- Algorithms and Data Structures, Complexity.
- Parallel and Distributed Computing
- Bioinformatics Applications; Biological Networks; Social Networks

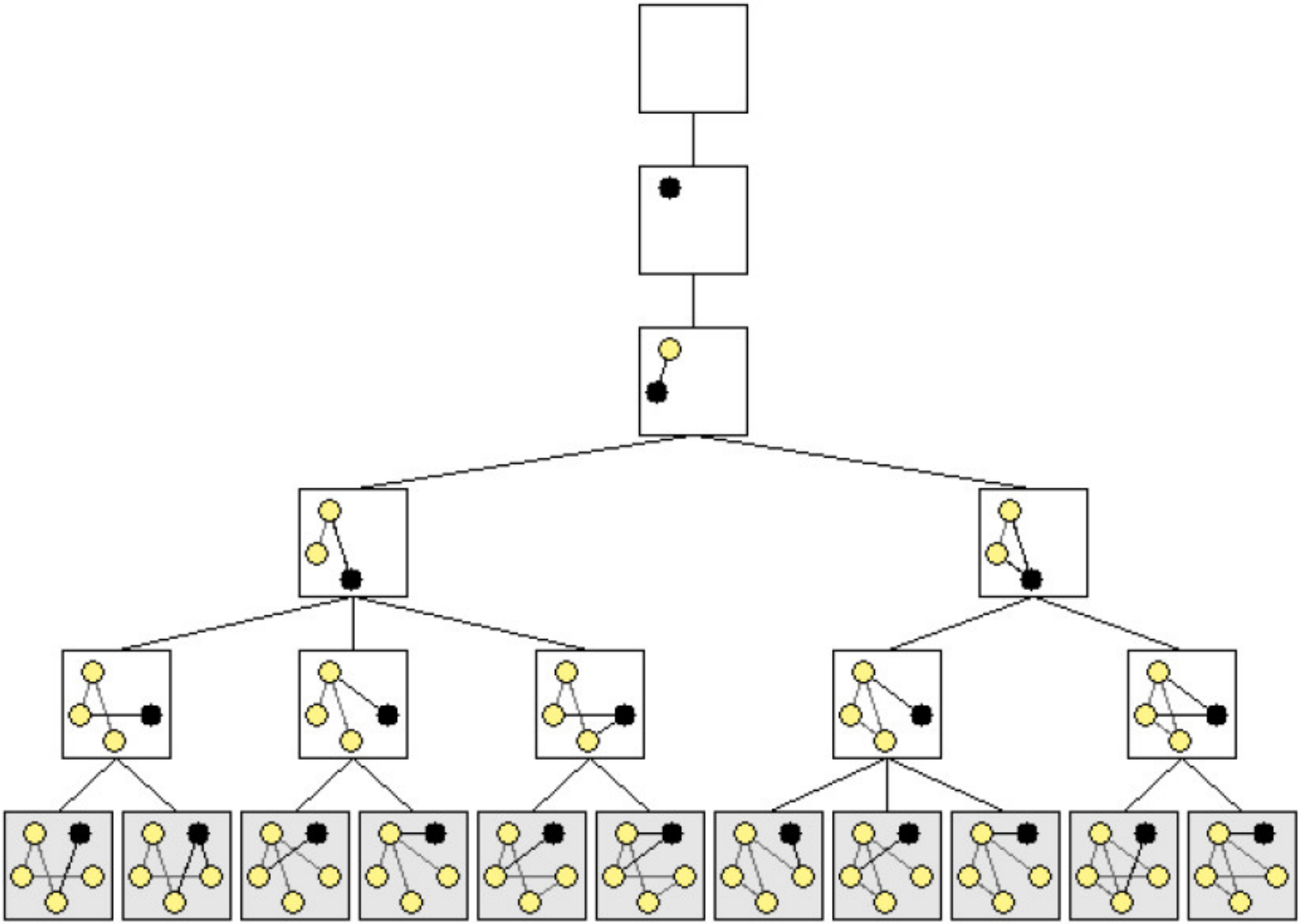
Other research interests:

- Computer Science Education and Programming Contests.
- Artificial Intelligence

PhD in Computer Science (2011)

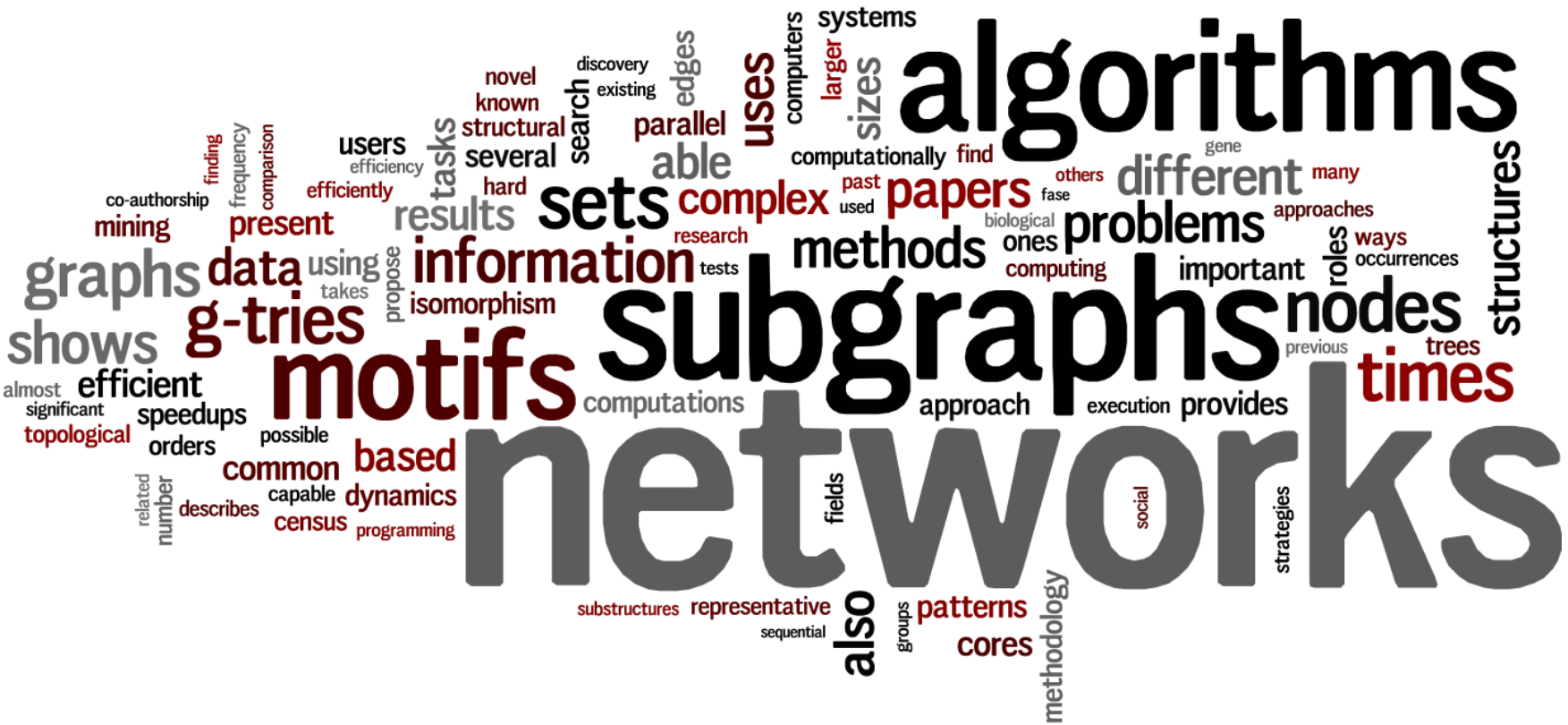
Thesis: Efficient and Scalable Algorithms for Network Motifs Discovery

PhD: G-Tries



Publications

A word cloud of my publications abstracts *(made in Dec 2015)*



http://www.dcc.fc.up.pt/~pribeiro/pubs_by_year.html

General Information

- **Site:** http://www.dcc.fc.up.pt/~pribeiro/aulas/arsi2021/
- **Classes:**
  - 1h30m (online - pre-recorded videos)
  - 30m (Zoom - time to be defined, non mandatory)
- **Class Organization:**
  - Theoretical Exposition: pre-recorded videos (YouTube)
  - Communication: Slack
- **Class participation:** There is no minimum enforced attendance

Evaluation

- Homeworks **(45%)**
- Presentation/Reviewing an Article **(15%)**
- Project **(40%)**

There will be no final exam.

There are no minimum grades in any of the evaluation components, but failure to deliver and present the project will result on a **RFC** evaluation (missing an evaluation component).

Homeworks

3 small individual homeworks (at least 3 weeks for each) to be delivered by email.

You will apply some concepts in practice (potentially using a computer to analyze small datasets).

You can discuss with me if you have difficulties when trying to do the homework

Predicted dates:

- **1st Homework (15%)** - somewhere in March
- **2nd Homework (15%)** somewhere in April
- **3rd Homework (15%)** - somewhere in May

Presentation/Reviewing an Article

- This is an **individual** assignment.
- Select a recent **scientific article** (year of publication ≥ 2016) about **network science**; carefully read it and present it to the class
- You should select a **topic that interests you** and confirm with the instructor that your choice is ok. Website will have some initial pointers to help in your search for a suitable article.

- **Date of presentation:** end of semester
- **Time for each presentation:** around 15 minutes

Project

- This is a **group** assignment (maximum: 2 students)
- **Analyze a network dataset** using the concepts you learned
- You can use your own built dataset of create a new one
- Focus can be on the **analysis itself** (ex: new insights), on the **implementation** (ex: compare efficiency) or any other combination
- You are strongly encouraged to **speak with your instructor to validate your project** idea before starting your work.

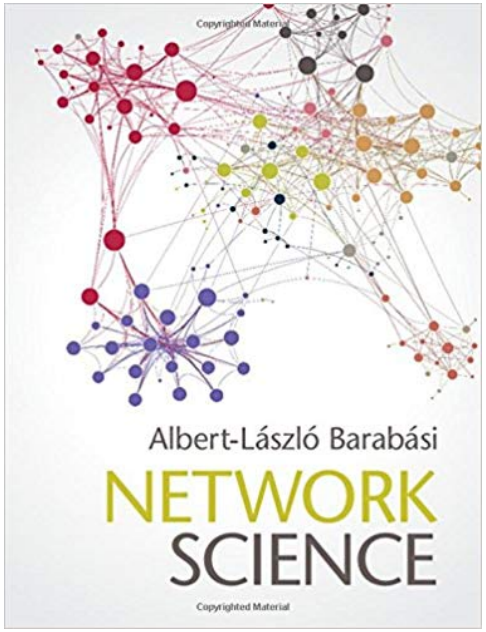
What do I need to deliver?

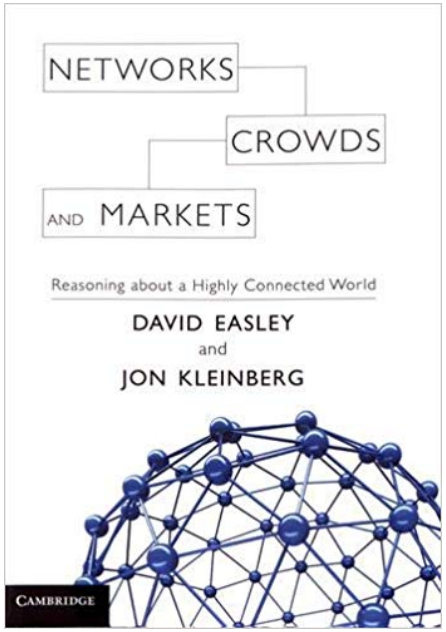
- **Written article:** 6 to 10 pages (KDD format)
  [deadline: end of classes]
- **Presentation:** 15 to 30 minutes

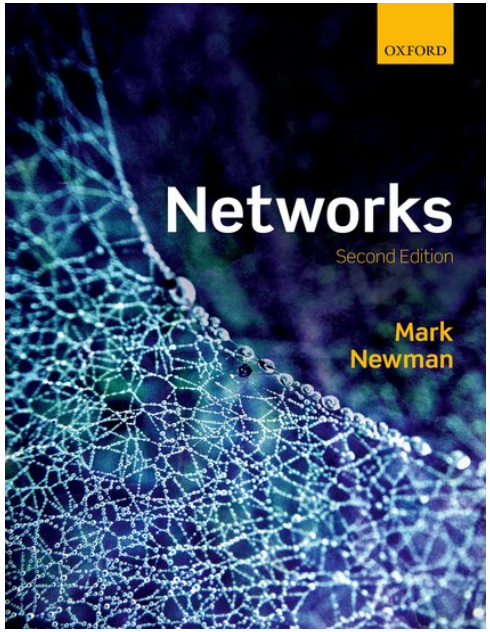
Learning outcomes in this curricular unit

- explain the **key concepts of network science** and network analysis
- apply a range of techniques for **characterizing network structure**
- define **methodologies for analyzing networks** of different fields
- demonstrate **knowledge of recent research** in the area

Recommended Books



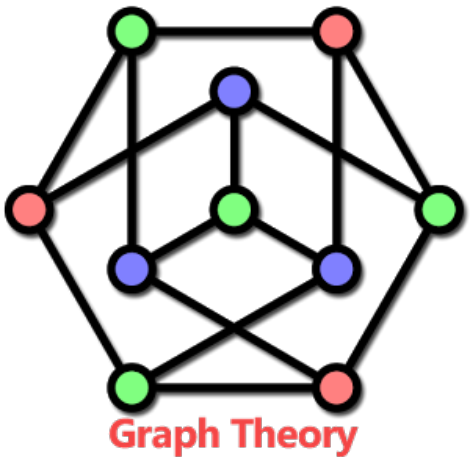




An overview of the program

Fundamentals of Network Science:

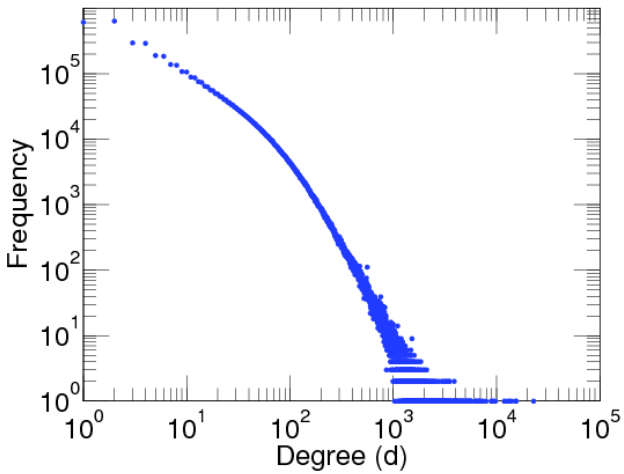
- Emergence of Network Science
- Essential Graph Theory
- Classic Graph Algorithms and Data Structures
- ...

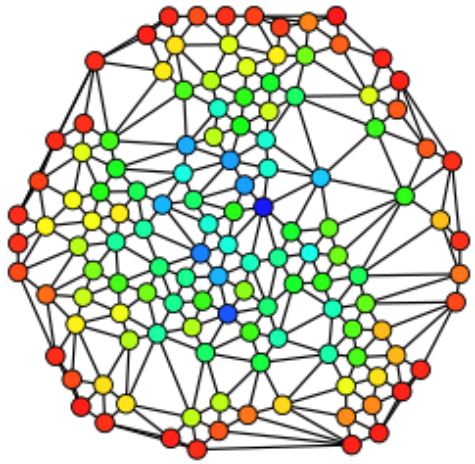


An overview of the program

Metrics and Basic Structural Properties:

- Degree distribution
- Paths and diameter
- Clustering coefficient
- Classic Node Centralities (betweenness, closeness, eigenvector, ...)
- ...

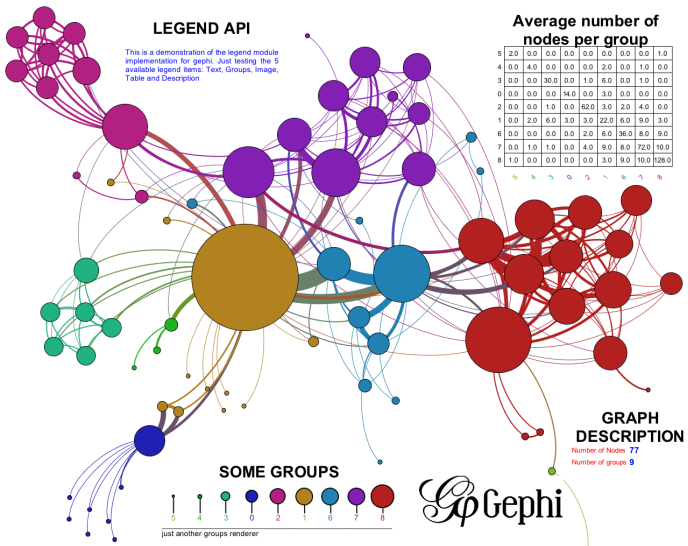


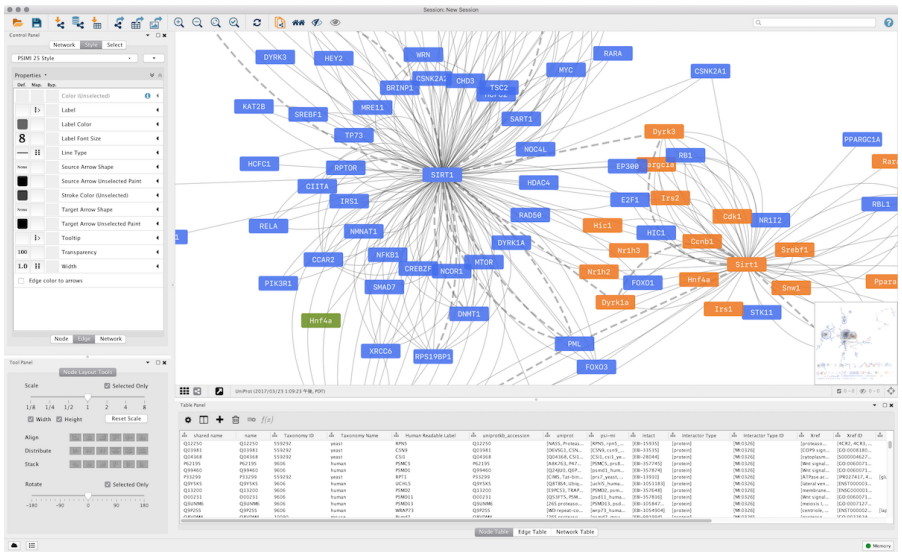


An overview of the program

Network Visualization:

- Graph Drawing
- Layout Algorithms
- Exploratory Analysis with Visualization
- ...

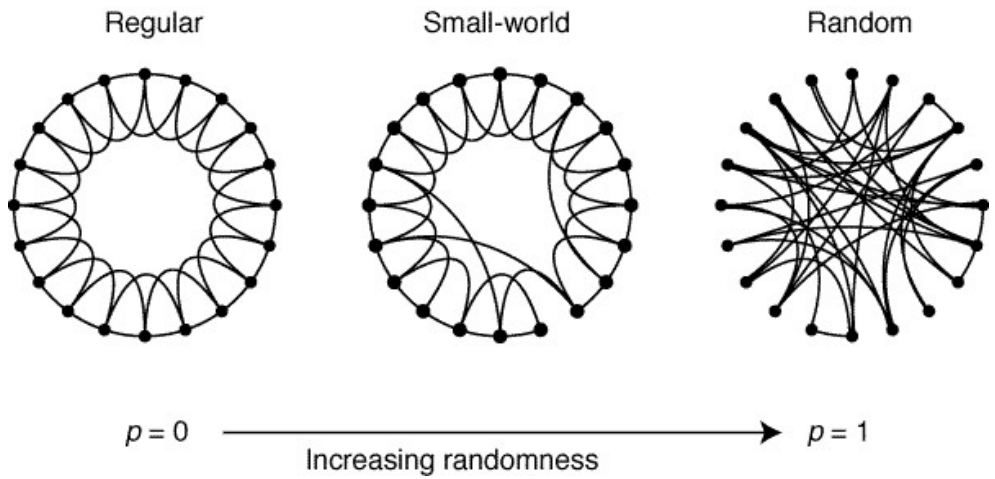




An overview of the program

Common Properties and Network Models:

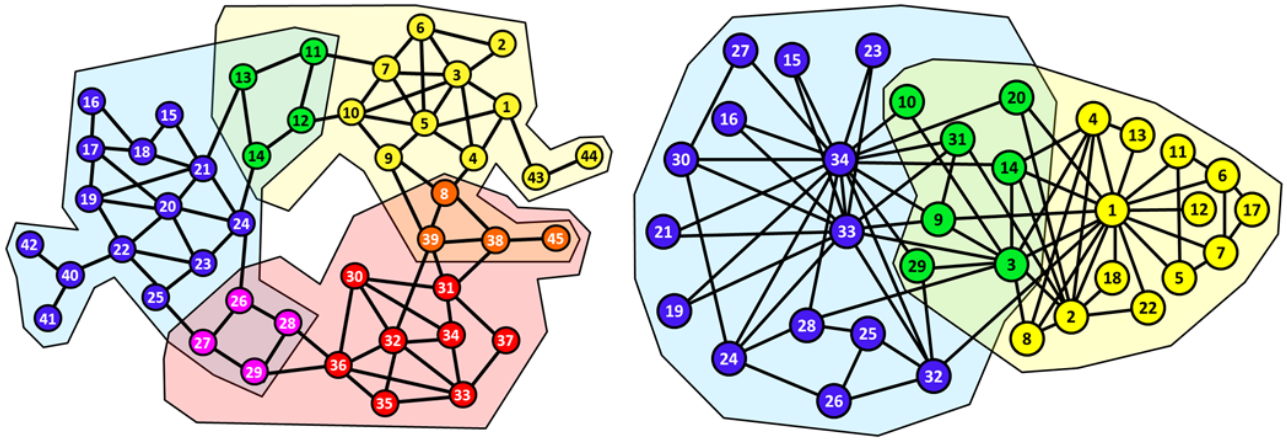
- Random networks and Erdös-Rényi model
- "small-world" property and Watts-Strogatz model
- "scale-free" property and Albert-Barabasi model
- Other models (ex: Kronecker graphs).
- ...



An overview of the program

Community Detection:

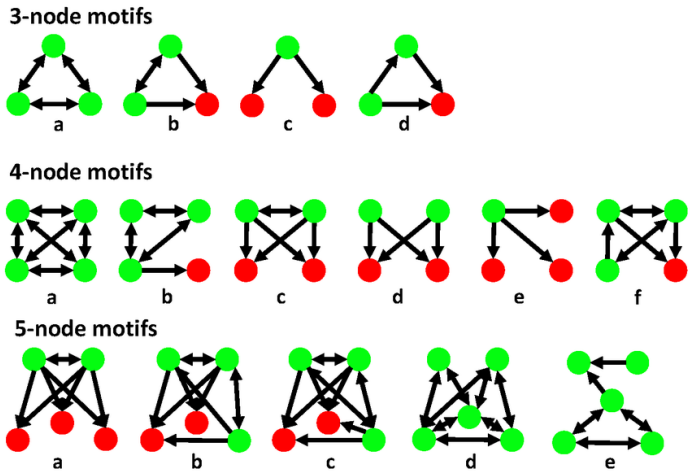
- Modularity
- Algorithms for detecting communities
- Overlapping communities and other variants
- ...

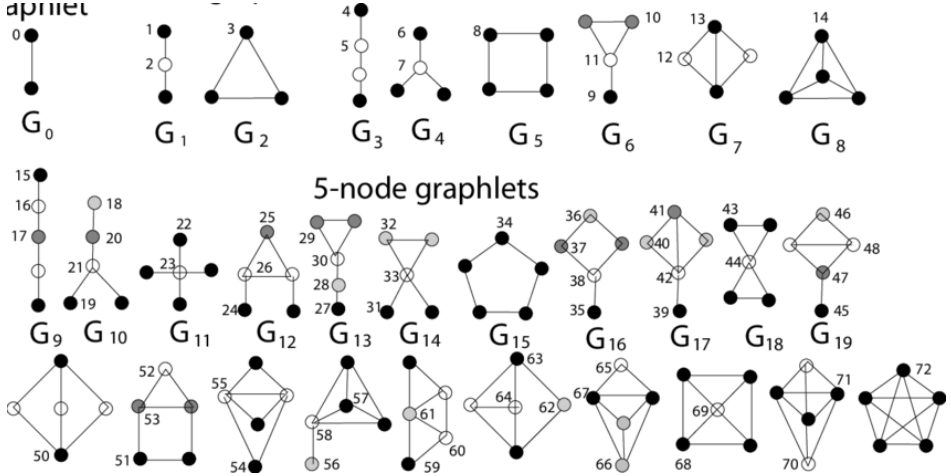


An overview of the program

Patterns and Subgraphs:

- Subgraph Census
- Network Motifs
- Graphlets
- Fingerprinting Networks
- ...

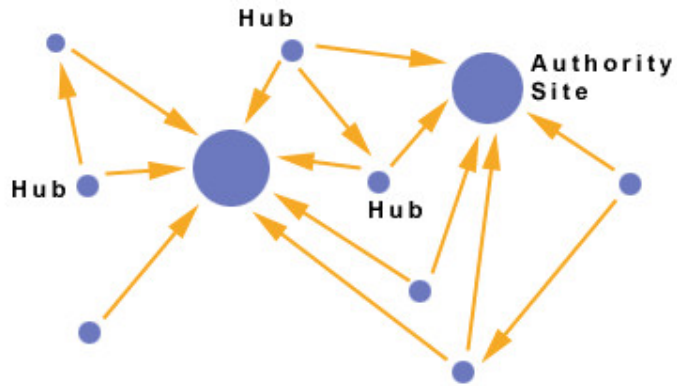




An overview of the program

Link Analysis:

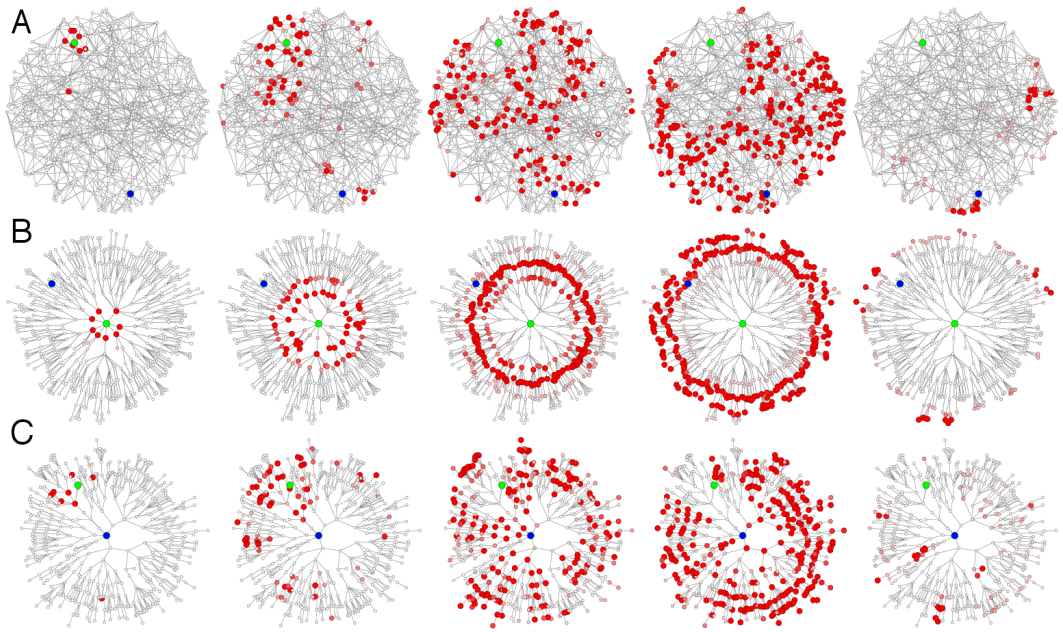
- Node Rankings
- HITS algorithm
- PageRank and other variants
- ...



An overview of the program

Propagation in networks:

- Information flow
- Epidemics
- Propagation models
- ...



An overview of the program

Other Selected Topics:

- Link Prediction
- Network Alignment
- Graph Databases
- Temporal Networks
- Multilayer networks
- Graph Neural Networks
- Graph APIs (networkx, igraph, snap, ...)
- ...