Title: Statistical analysis of network data

Brief description

This tutorial uses the igraph R package. It has two main parts, the first one a general introduction to networks and classic graph theory, and the second one about new(er) statistical methods. See the detailed outline below for details.

Everything is explained live, in R, via examples and a lot of different data sets. Attendees can follow the examples on their own machines, without having to type anything, and with the possibility to stop the tutorial any time, to inspect objects in their R session, or modify and rerun the tutorial code.

The tutorial strongly builds on our book with Eric Kolaczyk: Statistical Analysis of Network Data with R, and my book currently under construction (probably in press at the time of the conference): Network Analysis with R/igraph. (Co-authors Tamas Nepusz, and Edoardo Airoldi).

Detailed outline

Two parts, with a short break in between.

1. Why networks are interesting and important, via examples.

2. Creating and manipulating networks in R/igraph. Basic network concepts, classic graph theory. How to convert between different network data formats. How these tools can be used to implement new methods and algorithms quickly.

3. Paths and related concepts, searches, flows and cuts in networks. This is still classic graph theory and computer science.

4. Centrality in networks. Classic and modern measures, from degree/closeness/betweenness to PageRank.

5. Finding groups in networks, i.e. community structure detection. Various algorithms. Cohesive blocks.

6. Visualization of networks, layout algorithms. Visualization of large graphs, using results from community structure detection. Visualization with other R packages, or tools, like D3.js, Gephi, etc.

7. Random graphs models. Classic models, stochastic block models, mixed membership stochastic block models and their inference. Hierarchical random graphs.

8. Statistical methods: graph embeddings, graph matching, graphlets, graphons. Temporal networks and change-point detection. Connections to matrix factorizations.

Goals

1) Showcasing what can be done with R/igraph and other R packages. Give people a taste of the existing analysis methods, so that they can decide if and how they can use them for their own data sets.
2) Giving pointers to people for the various analysis methods. Attendees should be able to easily learn more about any mentioned method, by following the pointers and links in the tutorial.
3) Community building. Bringing together people interested in networks, so that they can exchange ideas and code, and can contribute to and benefit from R/igraph.

Importance

Analysis of networks is a popular area, both because of the abundance of network data, and the recent methodological developments.

Background knowledge required

Only basic R user skills, at the level of any decent introductory R textbook. Network concepts will be explained.

Potential attendees

Scholars, scientists, data analysts interested in network data methods, and/or how they can be implemented in R/igraph. People who have network data but do not know what to do with it (other than creating pretty spaghetti pictures). People who know what they want to do with their network data, but do not know how.