Introduction to Statistical Network Analysis Assignment

Lecturer: Dr. Riccardo Rastelli

Instructions

- The goal of this assignment is to analyse a network of social contacts at a conference, using the methods and tools that have been covered in the lecture.
- **Dataset.** The data is the network of face-to-face contacts of the attendees of the ACM Hypertext 2009 conference. The ACM Conference on Hypertext and Hypermedia 2009 (HT 2009, http://www.ht 2009.org/) was held in Turin, Italy over three days from June 29 to July 1, 2009. In the network, a node represents a conference visitor, and an edge means that the two individuals had a large number of contacts during the conference. The network is undirected.
- The undirected adjacency matrix is provided in the `RData` file. The full dataset and additional information are available from http://www.sociopatterns.org/datasets/hypertext-2009-dynamic-contact-network/
- You should write a report (up to 4 pages including tables and figures but excluding code) to show your analysis of this dataset. You should include the code used either through RMarkdown or in a separate file.
- You can find below a list of points that should be addressed in the report.

1 - Load the data, transform the adjacency matrix into a `igraph` object, plot the network.

2 - Calculate descriptive statistics of the network such as the number of nodes, number of edges, density, number of triangles, degrees and degree distribution.

3 - Perform spectral clustering. Comment on whether the algorithm can find a reasonable clustering structure. Comment on whether some of the groups exhibit community structure.

4 - Use the variational expectation-maximisation algorithm to fit a stochastic blockmodel to the data. Describe the connectivity behaviour for each of the groups found, discussing whether they exhibit community structure. Compare the optimal partition with the one obtained through spectral clustering.

FDSII Assignment: submission details

- The assignment and associated data will be emailed to you.
- Report must be a **pdf of no more than 4 pages**, including tables and figures but excluding code.
- You should submit the (commented) **code** used either through RMarkdown or in a separate file.
- Deadline: **5pm on 31st January 2022**.
- Email pdf + code to your institutional director i.e., James/Claire/Ken for UL/UCD/MU respectively.