

Project 3: Final Project

COMS 6998-006 Network Theory, Spring 2008

Topic Due: April 10, 2008
Entire Project Due: April 28, 2008

1 Overview

The final project is to contribute a significant project in network theory that could potentially be published. The project results should be documented in a 6-10 page Physics Review (PR) Style paper. Each student will be required to give an in class presentation on May 1, 2008.

2 Choosing A Topic

Due: April 10, 2008

You are required to have chosen a topic by Thursday, April 10. Send an email stating the following to Drago (radev@umich.edu) and Sara (ss3067@columbia.edu):

- Your name
- Title
- What sections of the class it is related to
- Background readings
- Description of the task
- Evaluation methodology (if applicable)

For example:

- Name: John Smith
- Title: Virus propagation over a small world network
- Section: Diffusion on graphs, Cascading behavior
- Readings:
 - D. Liben-Nowell, J. Kleinberg. Tracing Information Flow on a Global Scale Using Internet Chain-Letter Data. Proc. National Academy of Sciences, 105(12):46334638, 25 March 2008.
 - Email networks and the spread of computer viruses, M. E. J. Newman, Stephanie Forrest, and Justin Balthrop, Phys. Rev. E 66, 035101 (2002).
- Data: synthetically generated

- Description: I will generate a large number of small world networks. Then I will compare 5 different algorithms for virus propagation and determine to which ones the network is most susceptible.
- Evaluation: I will measure the speed of spread of the virus over the network as well as the status (infected, immune) of all nodes over time.

3 Final Project

Due: April 28, 2008,

Late Deadline: April 30, 2008 (10% penalty)

The final project consists of a 6-10 page PR-Style paper and all resources, such as any code and data.

The paper should include the following:

- Topic Description and motivation
- Related Work
- Datasets Used
- Evaluation
- Network statistics related to the topic
- Visualizations
- etc...

The final project can differ from student to student and different information may be required depending on the project. If you are uncertain about what else you should include (or exclude), discuss it with the Drago and Sara prior to submission.

4 Presentation

Each student will be required to give around a 10 minute presentation on their project in class on May 1, 2008.

5 Grading

The grading will be partially based on the originality of the work as well as the presentation. 5% will be based on peer review - everyone will have a sheet in which they will grade all presentations.

6 Resources

- A tutorial on how to run clairlib for this assignment is located at http://www1.cs.columbia.edu/~coms6998/Clairlib/clairlib_network_analysis.html.
- Software available on the clic machines, as well as instructions to run them, is listed at <http://www1.cs.columbia.edu/~coms6998/software.html> (Software: Pajek, Jung, Guess, NetworkX, Clairlib).
- Instructions to generate a PR Style paper are describe at http://www1.cs.columbia.edu/~coms6998/pr_style/instructions.html.

7 Submission

Your submission should include your dataset, any code you wrote, and your report.

You need to submit a TAR file with the following files:

- The raw data used to build the network.
- The network in Pajek and/or clairlib format.
- The PR-style paper (including latex + pdf + any images in separate files + makefile).
- Any software needed to preprocess or analyze the data.

Please name your files and directories using your last name and the homework number, e.g., John Smith will name his files and directories Smith1.

```
Smith3.tar:  
Smith3/paper  
Smith3/paper/Smith3.tex  
Smith3/paper/Smith3.pdf  
Smith3/data/Smith3.net  
Smith3/data/...  
Smith3/code/...  
Smith3/...
```