Homework 7

STAT-GB.4310: Statistics for Social Data
Instructor: Patrick O. Perry

Due April 19, 2016

1. Install the `ergm` and `latentnet` R packages. The `ergm` package includes Kapferer's network of interactions between 39 workers in a tailor shop in Zambia between June 1965 and August 1965. Load the dataset into R with the `data(kapferer)` command. Read the documentation for the dataset with `?kapferer`.

2. Produce a plot of the network using the `plot(kapferer)` command, using the default layout algorithm (`mode = "fruchtermanreingold"`).

3. Produce two more network plots with different layout algorithms. Specifically, try use the `plot` command twice, once with `mode="circle"` and once with `mode="kamadakawai"`. These layout algorithms are documented in `?plot.network` and `?network.layout`.

4. Use the `ergm` command (from the `ergm` package) to get an ERGM to the network, using terms for edges, 2-stars, and triangles. (Note: If the MCMC-CMLE fitting method fails, fit the model with pseudo-likelihood instead by specifying `estimate="MPLE"` in the `ergm` command.) Do the values of the fitted coefficients make sense to you? What, if anything, do you learn about the network from the ERGM fit?

5. Use the `ergmm` command (from the `latentnet` package) to fit a latent space network model with $\mathrm{d} = 2$ and $\mathrm{G} = 1$. Plot the resulting fit.

6. Fit latent space network models with $\mathrm{d} = 2$ and G ranging from 1 to 4. According to BIC, which model fits best? Is there evidence of clustering? If so, how many clusters are there?