Cathy Zhuang

QTM 385 Assignment 2

Dr. An

Question 1

In an article titled “Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook”, authors Andreas Wimmer and Kevin Lewis examine racial homophily and its relationships to endogenous ties. They took the Facebook profile pages of 1640 students and found that friends in pictures were the most adequate ties of study. Then, they proceeded to classify categories into facial categories and ethnic categories, with facial categories being White, Black, Asian, Hispanic, or Mixed, and with ethnic categories dependent on the user publicly signaling membership in an ethnic club. The main message of the study was that if one was to run ERGM, one would find strong racial and ethnic homophily. However, if one controlled for endogenous ties, the coefficient decreases by half. In other words, at least fifty percent of racial homophily could be explained by the endogenous ties.

I have a couple of questions about this study. First, I would like to see how racial homophily may differ within a racial or ethnic group, such as racial homophily between Asian-Americans versus international Asian students. Is there perhaps a different dynamic that goes into each, or do we see the same patterns of homophily even within racial or ethnic groups? I also had some questions about the use of pictures to identify ties. These days, it is common for you to be Facebook friends with someone you do not really know. For example, if someone gets a friend request from a person who is already Facebook friends with many of their other friends, they are likely to accept their friend request. It is also common for people to be tagged in pictures where they do not really know the other people in those pictures. Are pictures truly a good measure for ties in this case?

Question 2

Part 1

Table

Description automatically generated

Part 2

Table

Description automatically generated with medium confidence

Part 3

A picture containing text

Description automatically generated

Part 4

Text

Description automatically generated

From these results, it seems like model 2 has the smallest AIC, making it the best model, but not by much.

For the goodness of fit plots, only 3 plots showed up because I had to force MPLE. Dr. An said this was ok.

Chart

Description automatically generated

Chart, histogram

Description automatically generatedChart, histogram

Description automatically generated

The solid lines represent observed statistics - features of observed networks.

The dashed lines represent simulated statistics- the features of simulated networks.

That solid lines are within the 95% confidence intervals of the simulated lines indicates model convergence.

In the model statistics graph, the solid lines are on average not contained within the dashed lines or shaded areas. The same goes for our odegree and idegree graphs. This means that we are either overestimating or underestimating the values.

Thus, given that the solid lines are not entirely within the 95% confidence intervals of the simulated lines, our model did not converge really well.

Table

Description automatically generated

There are only a few statistically significant results in model 2. Of the statistically significant results: It seems like managers in department 3 are more likely to seek advice. The positive coefficient on gwidegree indicates that ties are less likely to concentrate on a few nodes. The negative coefficient on gwdsp indicates ties are less likely to form open triangles.

Of the not statistically significant results, we might say that: The negative coefficient on "edges" indicates that the network is sparser than expected. The positive coefficients on the age sender and receiver effects may mean that older managers are more likely to make and receive friendship nominations. Managers of the same age or same department are more likely to be friends. The odds for a tie to form a mutual relationship is e^0.3955 times the odds for the tie to form a non-mutual relation. The positive coefficient on GWESP indicates transitivity in tie formations; the odds for the tie to form a triangle is about e^0.003792 times the odds for the tie to not form a triangle. However, none of these results are significant, so we cannot say for sure.

R code:

library(statnet)

setwd("C:/Users/Cathy/Desktop/QTM 385 Network Analysis")

data <- as.matrix(read.csv("Manager\_advice.csv", header = T))

### Create a network object

Fnet <- network(data, directed = T)

### Import and set attributes for nodes

att <- read.csv("Manager\_att.csv", header = T, row.names = 1 )

set.vertex.attribute(Fnet, "age", att$age )

set.vertex.attribute( Fnet, "tenure", att$tenure )

set.vertex.attribute( Fnet, "level", att$level )

set.vertex.attribute( Fnet, "dept", att$dept )

model1 <- ergm(Fnet ~ edges + nodeicov("age") + nodeocov("age") + nodecov("tenure") +

nodefactor("dept") + absdiff("age") + absdiff("tenure") + nodematch("dept") )

summary(model1)

model2 <- ergm(Fnet ~ edges + nodeicov("age") + nodeocov("age") + nodecov("tenure") +

nodefactor("dept") + absdiff("age") + absdiff("tenure") + nodematch("dept")+

mutual + gwidegree(0.1, fixed = TRUE) + gwesp(0.1, fixed = TRUE) + gwdsp(0.1, fixed = TRUE), estimate = "MPLE",

control = control.ergm(MCMLE.maxit = 300, MCMC.burnin=10000, MCMC.interval=200, seed = 178 ))

summary(model2)

friends <- as.matrix(read.csv("Manager\_friends.csv", header = T))

anet <- network(friends, directed = T)

model3 <- ergm(Fnet ~ edges + nodeicov("age") + nodeocov("age") + nodecov("tenure") +

nodefactor("dept") + absdiff("age") + absdiff("tenure") + nodematch("dept")+

mutual + gwidegree(0.1, fixed = TRUE) + gwesp(0.1, fixed = TRUE) + gwdsp(0.1, fixed = TRUE) +

edgecov(anet), estimate = "MPLE", control = control.ergm(MCMLE.maxit = 300, MCMC.burnin=10000, MCMC.interval=200, seed = 178 ))

summary(model3)

AIC <- rbind(AIC(model1), AIC(model2), AIC(model3))

AIC

gof2 <- gof(model2)

par(mfrow=c(1,5))

plot(gof2)

summary(model2)