**Lab 9: Modeling Staff Communication Networks in R: ERGM**

We will continue our ERGM modeling of the clinic communication networks from Lab 8.

Complete all tasks and answer all questions for full credit.

**Task 1.** Lab 8 finished with a discussion of the necessary next steps to complete the modeling of the face-to-face communication network. Examine what a cycle term (ctriple) adds to the model.

setwd("Your Own Directory Here")

library(statnet)

F2FData <- read.table("PrimaryCareF2FNet.txt", header=TRUE, row.names=1, check.names=FALSE)

F2FMatrix <- as.matrix(F2FData)

StaffAttr<-read.table("PrimaryCareAttributes.txt",header=TRUE,

stringsAsFactors=FALSE)

F2FNet=network(F2FMatrix,matrix.type="adjacency",directed=TRUE)

F2FNet%v%'vertex.names'<- StaffAttr$SubjID

F2FNet%v%'Job\_Category'<- StaffAttr$Job\_Category

F2FNet%v%'Years\_Clinic'<- StaffAttr$Years\_Clinic

F2FNet%v%'FTE'<- StaffAttr$FTE

F2FNet%v%'Female'<- StaffAttr$Female

F2FNet%v%'Job\_Satisfaction' <- StaffAttr$Job\_Satisfaction

F2Fmodel.5 <- ergm(F2FNet ~edges + mutual + gwesp(0.25,fixed=TRUE) + ctriple + nodefactor("Job\_Category") + nodematch("Job\_Category",diff=TRUE)+ nodecov("Job\_Satisfaction") + absdiff("Years\_Clinic"), verbose=TRUE)

summary(F2Fmodel.5)

**Questions:**

**1. Was the ctriple term significant in the model? What does this say about clinic face-to-face communication ties?**

yes, it was significant. It means close triads increase density.

**Task 2.** Check goodness of fit for the new model.

gof(F2Fmodel.5, GOF = ~ distance + espartners + idegree + odegree + triadcensus, verbose=TRUE, burnin=1000, interval=1000)

**Questions:**

**2. Does adding the cycles term improve the model fit? In what ways?**

**yes. In the edgewise shared partner, there is less 0.00. It is doing better in triad census and geodesic distance as well. We are doing better in high end.**

**//significat negative: we did not even do good as random. In case of triad, negative triad means triads are not even randomly closed.**

**Task 3.** We now turn our attention to electronic communication in the clinic. We will repeat many of the steps used to model the face-to-face communication. Read in the EHR communication network data.

EHRData <- read.table("PrimaryCareEHRNet.txt", header=TRUE, row.names=1, check.names=FALSE)

EHRMatrix <- as.matrix(EHRData)

StaffAttr<-read.table("PrimaryCareAttributes.txt",header=TRUE,

stringsAsFactors=FALSE)

EHRNet=network(EHRMatrix,matrix.type="adjacency",directed=TRUE)

EHRNet%v%'vertex.names'<- StaffAttr$SubjID

EHRNet%v%'Job\_Category'<- StaffAttr$Job\_Category

EHRNet%v%'Years\_Clinic'<- StaffAttr$Years\_Clinic

EHRNet%v%'FTE'<- StaffAttr$FTE

EHRNet%v%'Female'<- StaffAttr$Female

EHRNet%v%'Job\_Satisfaction' <- StaffAttr$Job\_Satisfaction

Check the density, transitivity, and centralization of the EHR network.

gden(EHRNet)

gtrans(EHRNet)

centralization(EHRNet, betweenness)

Plot the EHR communication network.

EHRBetw <- (betweenness(EHRNet)/max(betweenness(EHRNet))\*2)+0.5

Jobcat<- EHRNet%v%'Job\_Category'

gplot(EHRNet, vertex.cex= EHRBetw,vertex.col=Jobcat,usearrows=TRUE)

legend("bottomleft", legend = c("MD","RN","MA","Office","Lab"),fill = c(1:5), cex = 0.8)

Check the mixing matrix of the EHR communication.

table (Jobcat)

mixingmatrix (EHRNet, "Job\_Category")

**Questions:**

**3. How do the graph and the mixing matrix for EHR communication network differ from the face-to-face communication network?**

**very high from nurse(2), office(4) but very low from medical assistance(3). It has more inter-job category communication.**

**For F2F, NURSES have very higher connection.**

**Task 4.** Begin the ERGM modeling

EHRmodel.0 <- ergm(EHRNet ~edges, verbose=TRUE)

summary(EHRmodel.0)

EHRmodel.1 <- ergm(EHRNet ~edges + nodefactor("Job\_Category"), verbose=TRUE)

summary(EHRmodel.1)

**Questions:**

**4. Which categories of employees are more likely to send and receive communication ties? Which are least likely? Why do you think this is?**

**2(nurse) - more likely to send or receive ties**

**3(MA), 5(lab) - less (neg sig) likely to send or receive ties**

**4 - same as phycisians, p value is not significant**

//estimate - log(odd ratio)

**Task 5**. Try adding nodematch.

EHRmodel.2 <- ergm(EHRNet ~edges + nodefactor("Job\_Category") + nodematch("Job\_Category",diff=TRUE), verbose=TRUE)

summary(EHRmodel.2)

**Questions:**

**5. What happened when the nodematch term was added? Why?**

it does not have much significant predictive power. but actually, we can not trust this model as it has a degenerate model (nodematch part) because there was a -inf estimate(due to a 0 in mixingmatrix).

**Task 6**. Add job satisfaction and number of years in the clinic. Remove nodematch.

EHRmodel.3 <- ergm(EHRNet ~edges + nodefactor("Job\_Category") + nodecov("Job\_Satisfaction") + absdiff("Years\_Clinic"), verbose=TRUE)

summary(EHRmodel.3)

**Questions:**

**6. Are EHR network connections related to job satisfaction or number of years in the clinic? Why you might this to be the case?**

yes. less job satisfaction means more electronic connections. also, less years in clinica means more electronic connection. I think maybe more electronic connection and less f2f is making people less satisfied in job.

//predicting the correlation between two factors, we do not the causation.

**Task 7.** Add structural terms to the model.

EHRmodel.4 <- ergm(EHRNet ~edges + mutual + ctriple + esp(1) + nodefactor("Job\_Category") + nodecov("Job\_Satisfaction") + absdiff("Years\_Clinic"), verbose=TRUE)

summary(EHRmodel.4)

**Questions:**

**7. Are the structural terms significant in the model?**

**Yes. Mutual and edge-wise shared partner are significant.**

**esp1 -> do you have at least one shared partner? much more likely to have more than one in this case.**

**Task 8**. Summary. Summarize your findings about network connections in a clinic setting. Discuss additional hypotheses you would interested in testing with regard to clinic communication structure.

From the analysis, we see that it has more inter-job category connections. That’s why node match did not have much predictive power. However, we see that job satisfaction and number of years have significant predictive power as well. It would be interesting to see if less job satisfaction is creating more electronic communication or more electronic communication is creating less job satisfaction. Though testing causality might be complex.