Network Analysis: Homework

Jeff Ziegler

Due: August 23, 2018

1 Nigeria Data Processing

a) Process the data: turn this event dataset into a matrix.

b) Specifically, summarize the interactions across all time periods into an adjacency matrix where:

1. "1" indicates that i and j had a conflictual interaction sometime during the temporal span of the original dataset and zero otherwise.
2. Make sure all actors that existed at any point during the temporal span are included in the adjacency matrix.

```
rm(list=ls())
# set working directory to Git location
setwd('/Users/jeffziegler/Documents/Git/')
# load libraries
#slibrary(tidyverse); library(igraph); library(plyr)
# load data
load("network2018_hw1/nigeria.rda")

################################
# (1) Nigeria Data Processing
################################

# create variable for row and column length of adjacency matrix to fill
rowLength <- length(unique(nigeria$sender))
colLength <- length(unique(nigeria$receiver))
# create adjacency matrix of sender and receiver, filled w/ zeroes
nigeriaAdjMat <- matrix(0, nrow=rowLength, ncol=colLength)
# adjust row and column names for sender and receiver
rownames(nigeriaAdjMat) <- unique(nigeria$sender)
colnames(nigeriaAdjMat) <- unique(nigeria$receiver)
# fill in adjacency matrix per year (i)
# start by sorting all unique years to iterate over
nigeriaAdjMatYearlyList  <- lapply(sort(unique(nigeria$year)), function(i){
  # find just those pairings for (1) a given year
```

```
  currentYear <- nigeria[nigeria$year == i,]
  # and (2) had a conflict (conflict == 1)
  yearlyConflicts <- currentYear[currentYear$conflict == 1,]
  # now that we know which pairings had conflicts
  # fill in a "1" based on sender and receiver
  for(i in 1:nrow(yearlyConflicts)){
    nigeriaAdjMat[as.character(yearlyConflicts[i,]$sender),
                  as.character(yearlyConflicts[i,]$receiver)] <- 1
  }
  # return the adjacency matrix, which will be placed in a list
  return(nigeriaAdjMat)
})
# collapse all the matrices in list into one matrix
# since instructions are to "summarize the interactions
# across all time periods into a single matrix"
nigeriaAdjMatTotalList <- Reduce('+', nigeriaAdjMatYearlyList)
```

2 Measurements & Community Detection

a) Which actor is the most "influential" in the network? Justify your response and the measure you choose to estimate "influence."

b) Employ the blockmodel function from the sna package to explore potential group level structure in the data (see slides 61-63 from day 2 for details):

- Run blockmodel with varying levels of k.
- Save the node classifications from each run.
- Now how do we choose k?
  - You will do so through an out-of-sample cross-validation exercise (at least 10 folds).
  - Report the AUC (ROC) and AUC (PR) statistics from each model.

c) After having determined the k that gives the best out of sample performance, visualize your results as shown in slide 67 from the day 2 lecture

```
##############################################
# (2) Measurements and Community Detection
##############################################

# (a) We want to discern who the most "influential"
# actor in the network is.
# We'll create two measures of "influence":
# (1) degree (number of connections)
# (2) eigenvector centrality (connections to high-scoring nodes
# contribute more to the score of the node)

# create graph object from igraph package
```

```r
library(igraph)
nigeriaGraph <- graph_from_adjacency_matrix(nigeriaAdjMatTotalList,
          mode='directed', weighted=TRUE, diag=FALSE)

# (1) degree
head(sort(degree(nigeriaGraph), decreasing=T))

# interestingly, the Police (Nigeria) and the Military (Nigeria)
# are two of the top 3 most engaged actors (Fulani Militia is #2)

# (2) eigenvector centrality (connections to high-scoring nodes
head(sort(eigen_centrality(nigeriaGraph, directed = TRUE)$vector,
          decreasing=T))

# again, the police and military are not only more involved in conflicts
# but they engage w/ other highly conflicted actors
```

3 ERGMs

a) Run a cross-sectional ERGM on the Nigerian conflict network, develop at least one or two network level hypotheses.

b) Briefly discuss the results.

c) Make sure to show that you checked for convergence.

4 Find your own data

a) Locate data that relates to your field of interest.

b) Transform the data, or a subset of it into a matrix, and plot (similar to step 1 in Section 1).

c) Include descriptive features in your network graph (similar to step 2, but choose your own measurements).

d) Run a model, it can be any network model from the course but justify your choices!

e) Discuss the results in a brief write up. Present for 3-5 minutes to the class.