Exponential Random Graph Models
Statistical Modelling of Political Networks

Marc Schalberger

July 6, 2022

Outline

ERGM

- ➢ Idea
- ➢ Model
- ➢ Estimation
- ➢ Limitations

Application

- ➢ Data
- ➢ ERGM
- ➢ GOF

Exponential Random Graph Model (ERGM)

ERGM

Exponential Family Random Graph Model (ERGM) is one of the most popular approaches in statistical network analysis

Also known as p* models

First introduced by Ove Frank and David Strauss in 1986 in connection with the paper "Markov Graphs"

Widespread also in non-statistical circles such as social or political studies

ERGM Idea

Basic objective: construct a probability distribution which can be used generate network

Idea: Rather than the ties being the dependent variable, the network as a whole is the dependent variable
- ➢ model tries to estimate what the most likely network configuration is

The independent variables can be exogenous or endonegnous effects

- ➢ Exogenous effects are those occuring outside the outcome network (e.g. age, wage)

- ➢ Endogenous effects are those occuring within the outcome network (e.g. number of ties, reciprocity)

ERGM Model

The probability of observing network N over all possible networks we could have observed

$$P(N, \theta) = \frac{exp\{\theta' h(N)\}}{\underbrace{\sum_{N^* \epsilon \mathcal{N}} \exp\{\theta' h(N^*)\}}_{k}}$$

$\mathcal{N}$ set of all possible networks

θ vector of parameters

$h(N)$ vector of network statistics (exogenous or endogenous)

N^* element of $\mathcal{N}$

k normalizing constant

ERGM Assumptions

Assumptions:

1. equal probability of observing any two networks with the same values for the vector of statistics h
   ➢ no omitted variables

2. given a set of observed network statistics h, the observed network exhibits the average value of those statistics over the networks that could have been observed

ERGM Specification

Endogenous Dependencies:

Total Number of Ties $h(N) = \sum_{i<j} N_{ij}$

Reciprocity $h(N) = \sum_{i<j} N_{ij}N_{ji}$

Transitivity $h(N) = \sum_{i<j<k} N_{ij}N_{jk}N_{ik}$

Exogenous Dependencies:

Node Factor $h(N,X) = \sum_{ij} N_{ij}X_{ij}$

ERGM Interpretation

Edge-Level Interpretation:

Odds
$$\text{odds}\left(N_{ij} = 1 \middle| N_{ij}^c\right) = \frac{\exp\{\theta' h(N^+)\}/k}{\exp\{\theta' h(N^-)\}/k} = \exp\{\theta' \delta^{(ij)}(N)\}$$

$$\delta^{(ij)} = h(N^+) - h(N^-)$$

Log-Odds
$$logit\left(N_{ij} = 1 \middle| N_{ij}^c\right) = \log\left(\frac{\exp\{\theta' h(N^+)\}/k}{\exp\{\theta' h(N^-)\}/k}\right) = \theta' \delta^{(ij)}(N)$$

$$\delta^{(ij)} = h(N^+) - h(N^-)$$

ERGM Estimation

Estimation of θ is difficult because the likelihood is intractable

Markov Chain Monte Carlo (MCMC)

Default for most software packages

Normalisation constant k is approximated using a series of sample networks

Maximum Pseudolikelihood

Analytic approximation

Maximum pseudolikelihood computed by using hill-climbing algorithm to find parameters that maximize

$$\log(\prod_{i=2}^{n}\prod_{j=1}^{i} p_{ij}^{N_{ij}}(1-p_{ij})^{1-N_{ij}})$$

ERGM Limitations

Limitations:

Tendency to be numerically unstable in estimation
➤ Degeneracy

Deciding which endogenous effects to include in the model requires strong theoretical knowledge

ERGM cannot model longitudinally observed networks
➤ Temporal ERGM (TERGM)

ERGM unable to accommodate weighted networks
➤ Valued ERGM

Application

Dataset

"Networks of Violence: Predicting Conflict in Nigeria" by Cassy Dorff, Max Gallop and Shahryar Minhas (2020)

Raw data comes from the Armed Conflict Location and Event Data Project (ACLED)

Nodes: societal groups in Nigeria

Ties: armed conflict between actors in a given year (non-directed)

Dataset

Nigerian Intra–State Conflict
(2000–2016)

Dataset

2001 (Pre Boko Haram) | 2002 (Pre Boko Haram) | 2003 (Pre Boko Haram) | 2004 (Pre Boko Haram) | 2005 (Pre Boko Haram) | 2006 (Pre Boko Haram) | 2007 (Pre Boko Haram) | 2008 (Pre Boko Haram)

2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016
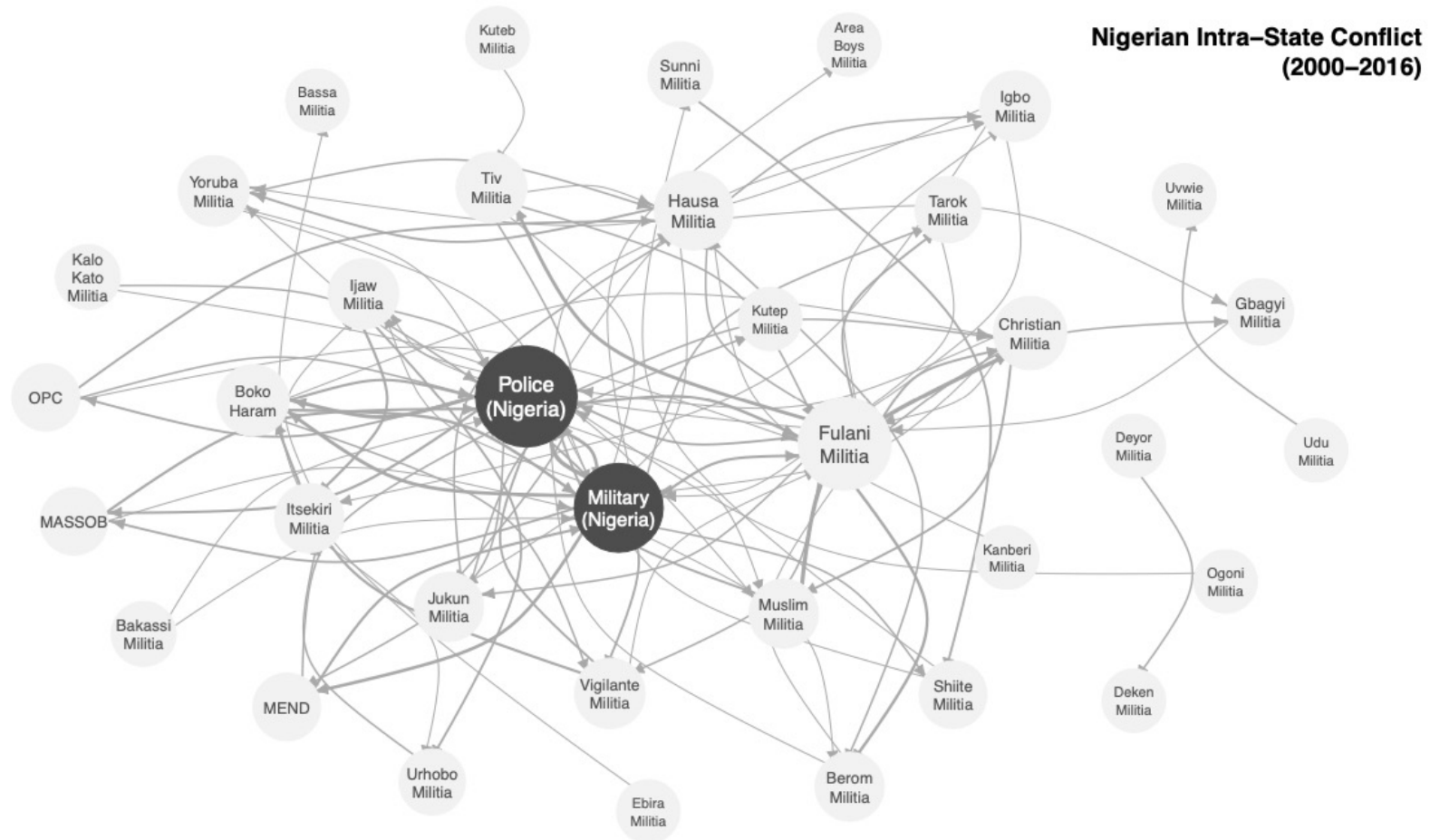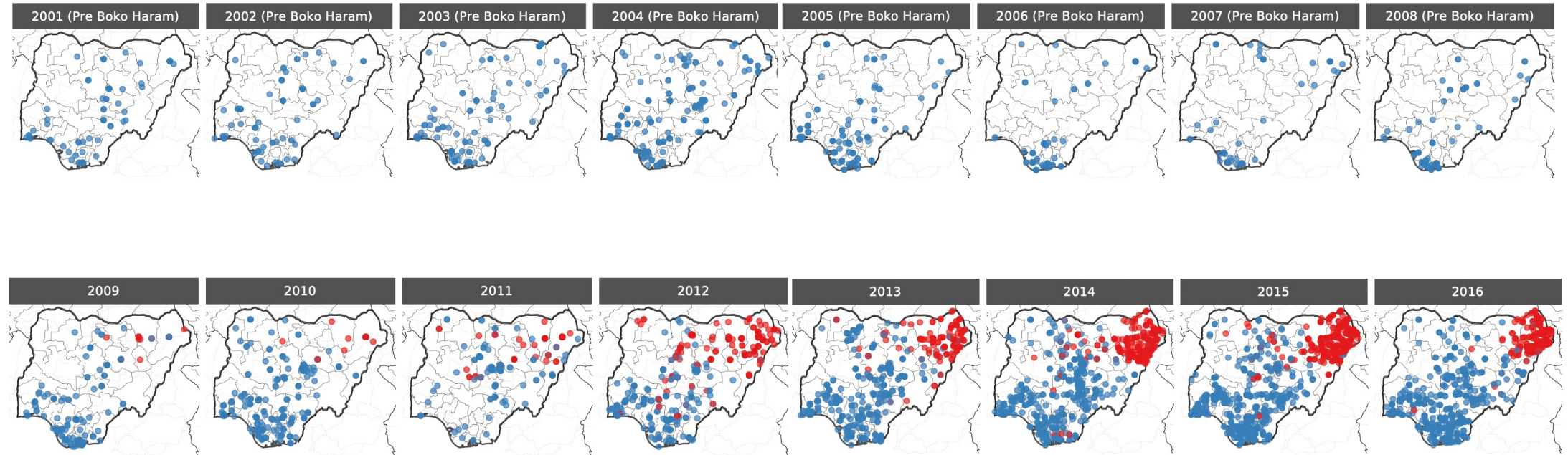
Confict Involving Boko Haram?   ● Yes   ● No

Dataset
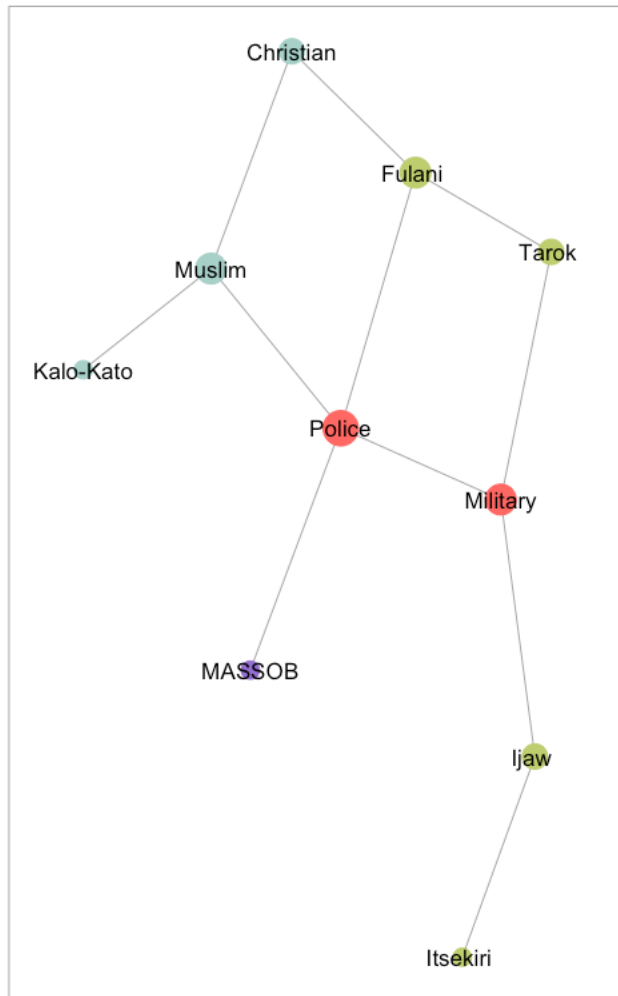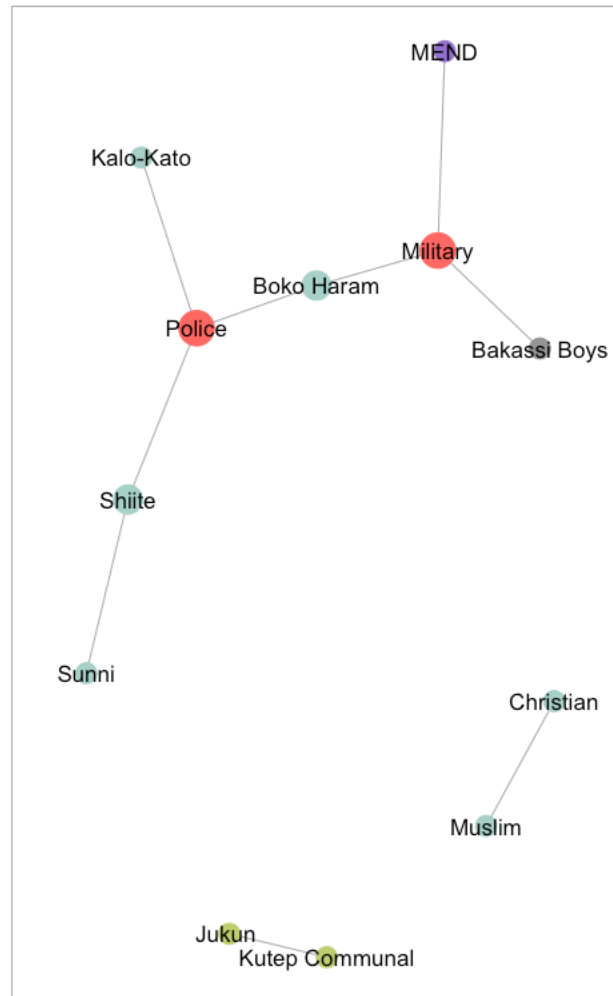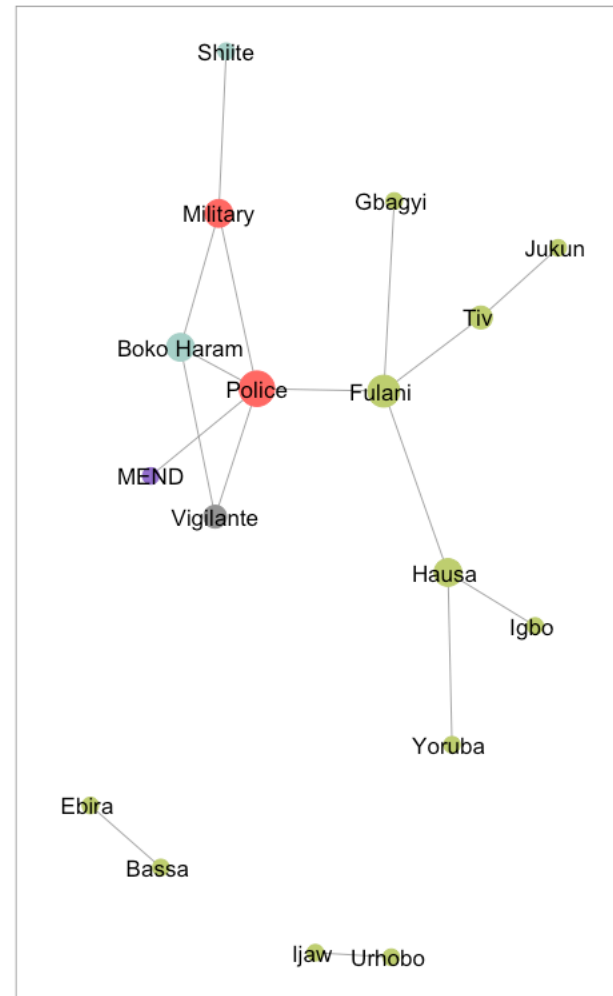
ERGM Specification

Endogenous Effects:

- Total number of ties

- Geometrically-weighted degrees

Exogenous Effects:

- Violent events against civilians

- Actor categories

- Riots/Protests

- Geographic spread

ERGM Results

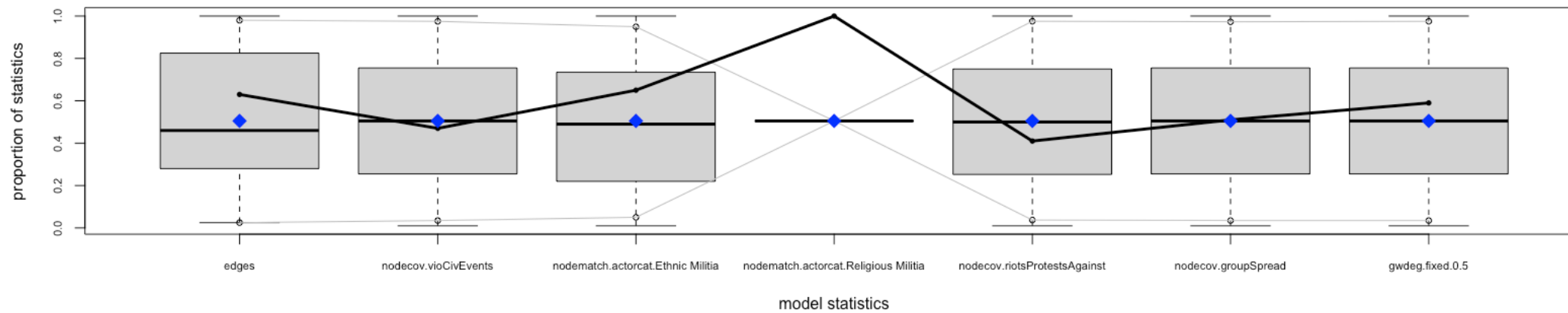|	2004	2009	2016
Covariates			
Number of Edges	-5.22 (0.001) **	-7.48 (0.005) **	-13.96 (0.0006) ***
G-W Degrees	-1.70 (0.03) *	0.76 (0.58)	4.32 (0.02) *
Violence Against Civ.	0.73 (0.01) *	0.23 (0.24)	0.04 (0.004) **
Match Ethnic	1.92 (0.12)	0.64 (0.64)	7.29 (0.003) **
Match Religious	4.88 (0.0004) ***	4.50 (0.002) **	NA
Riots/Protests	-1.40 (0.02) *	NA	0.05 (0.12)
Geo Spread	1.11 (0.01) *	1.08 (0.04) *	2.43 (0.03) *

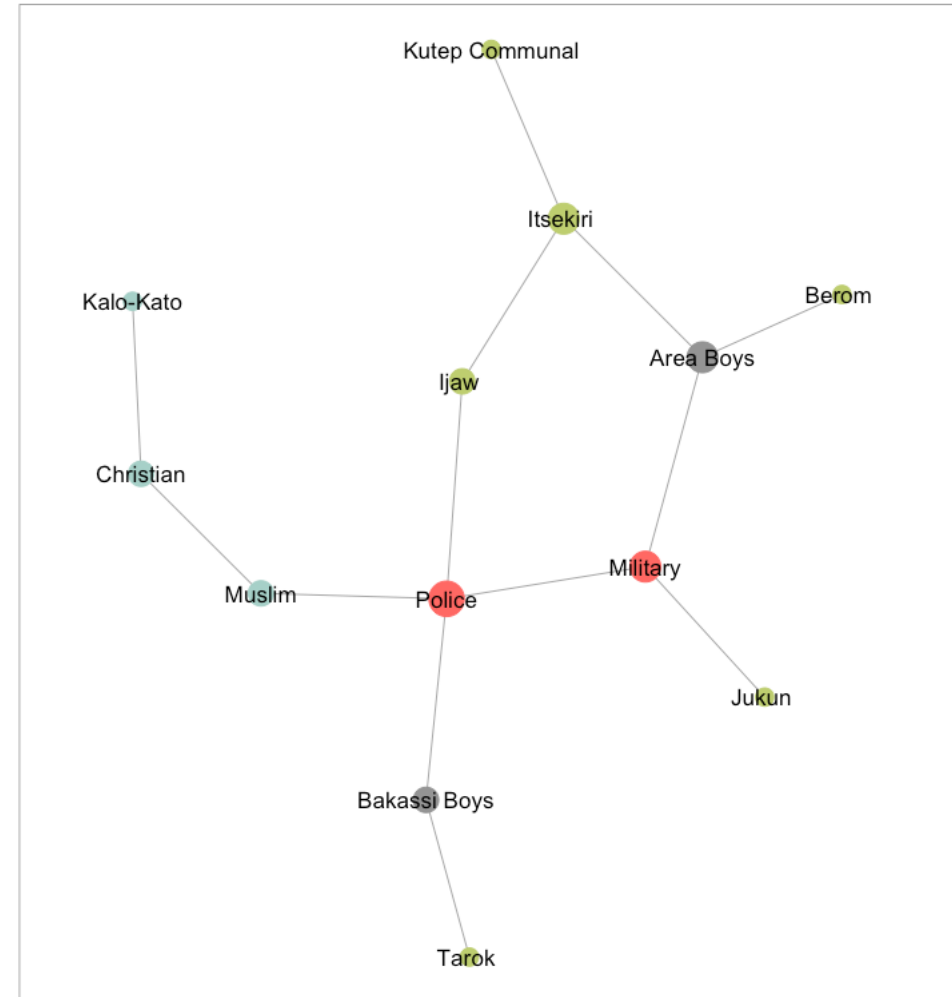# ERGM GoF 2004
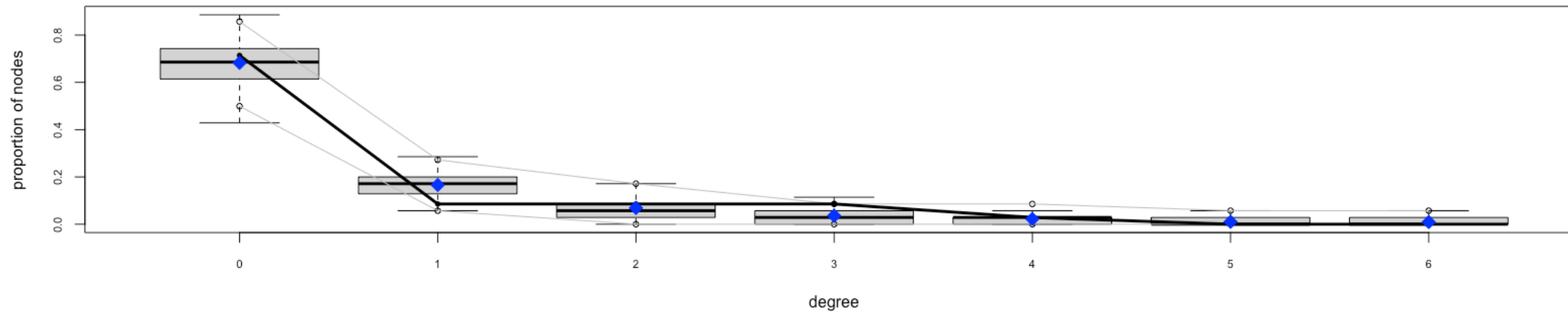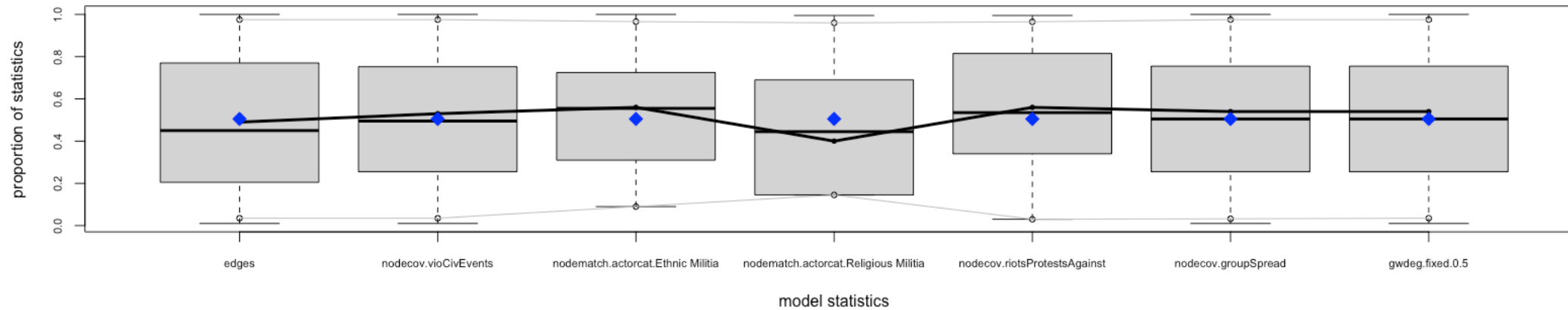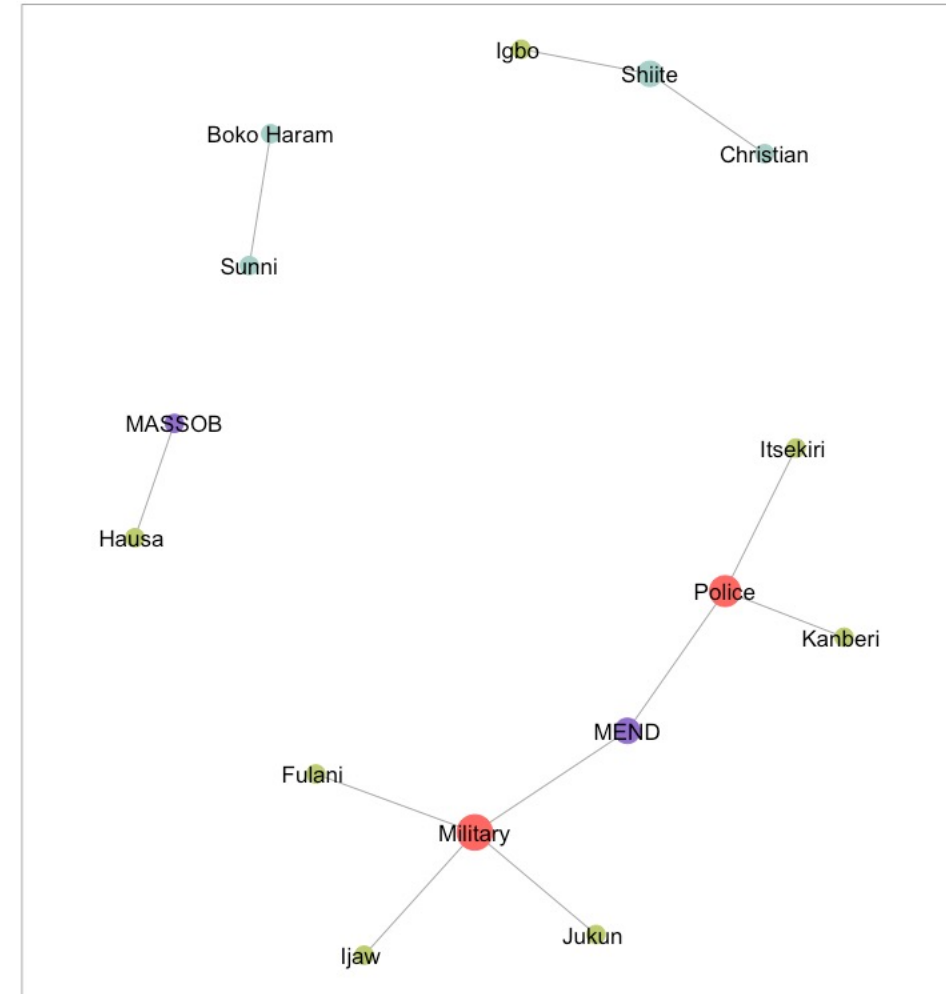
ERGM GoF 2009
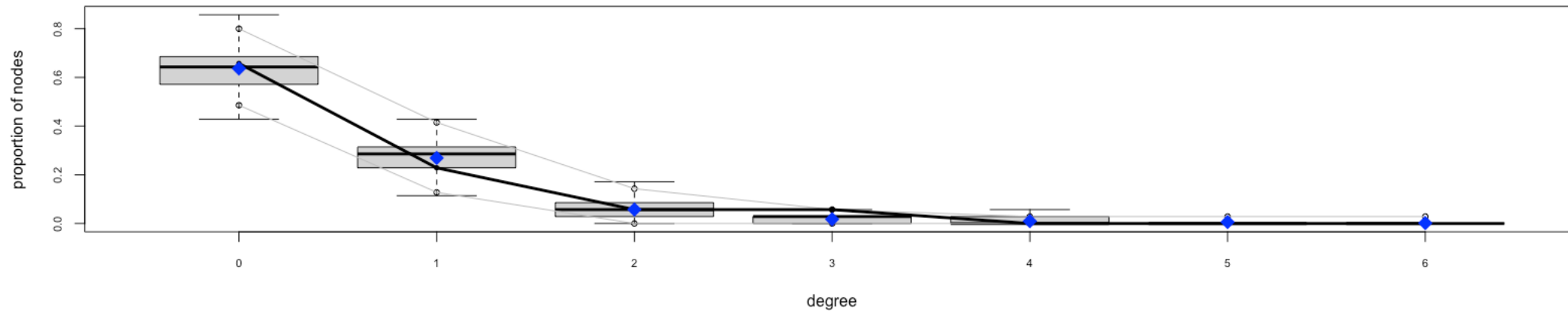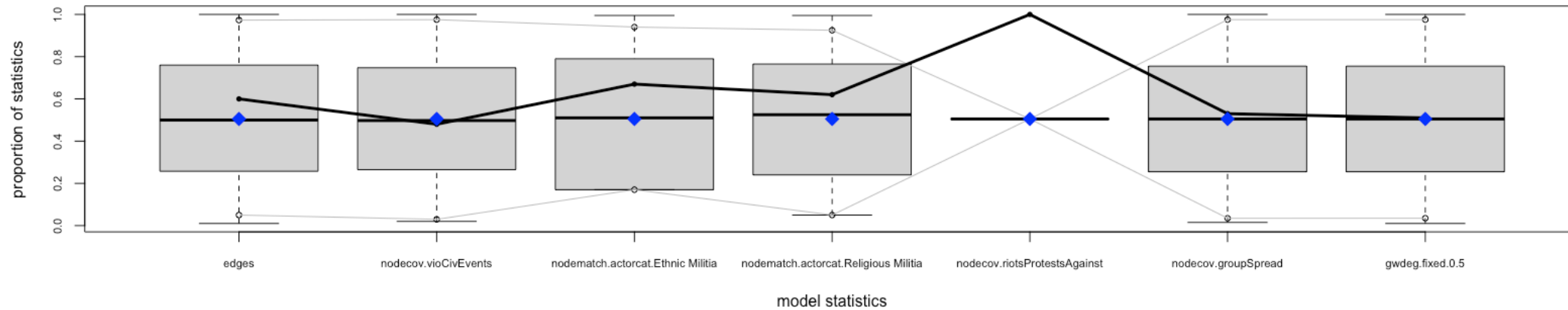
ERGM GoF 2016
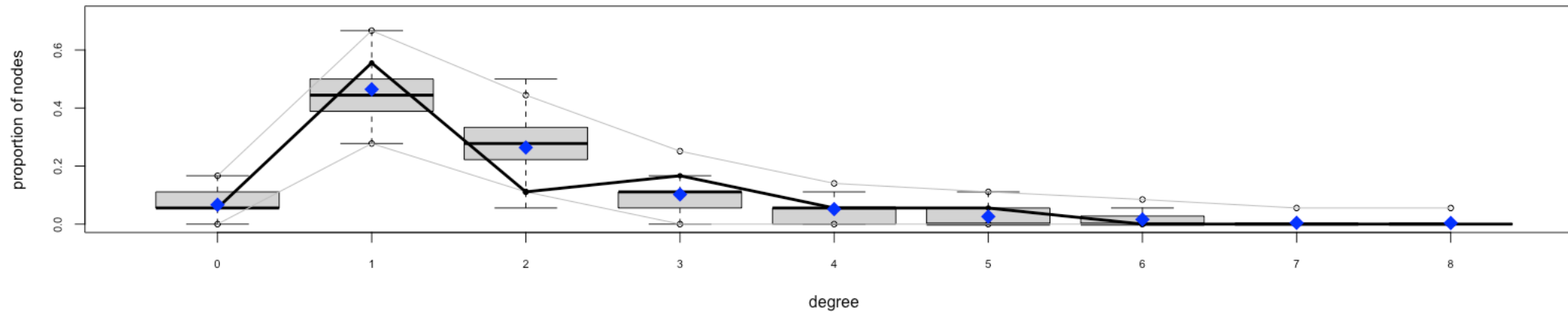
Simulation 2004

Simulation 2009

2009
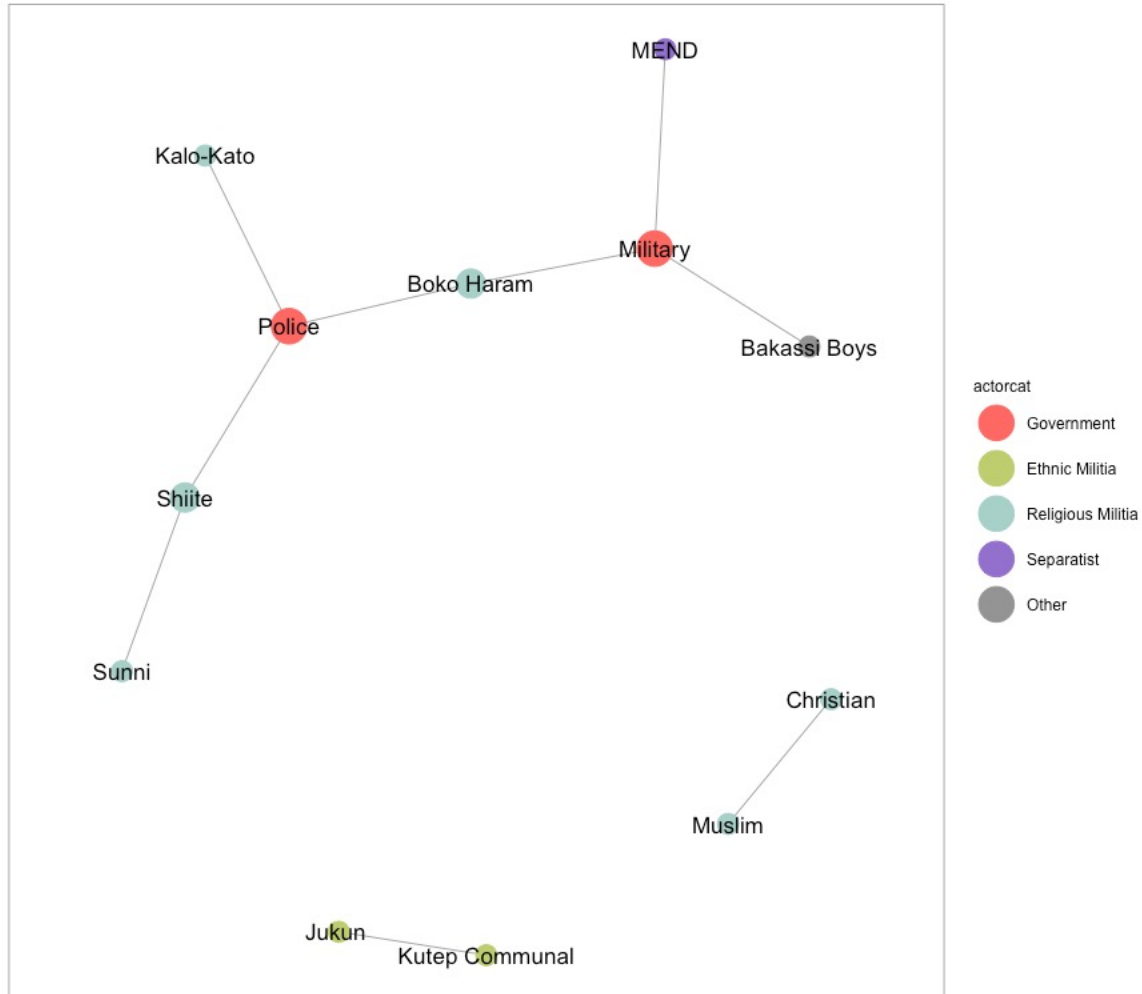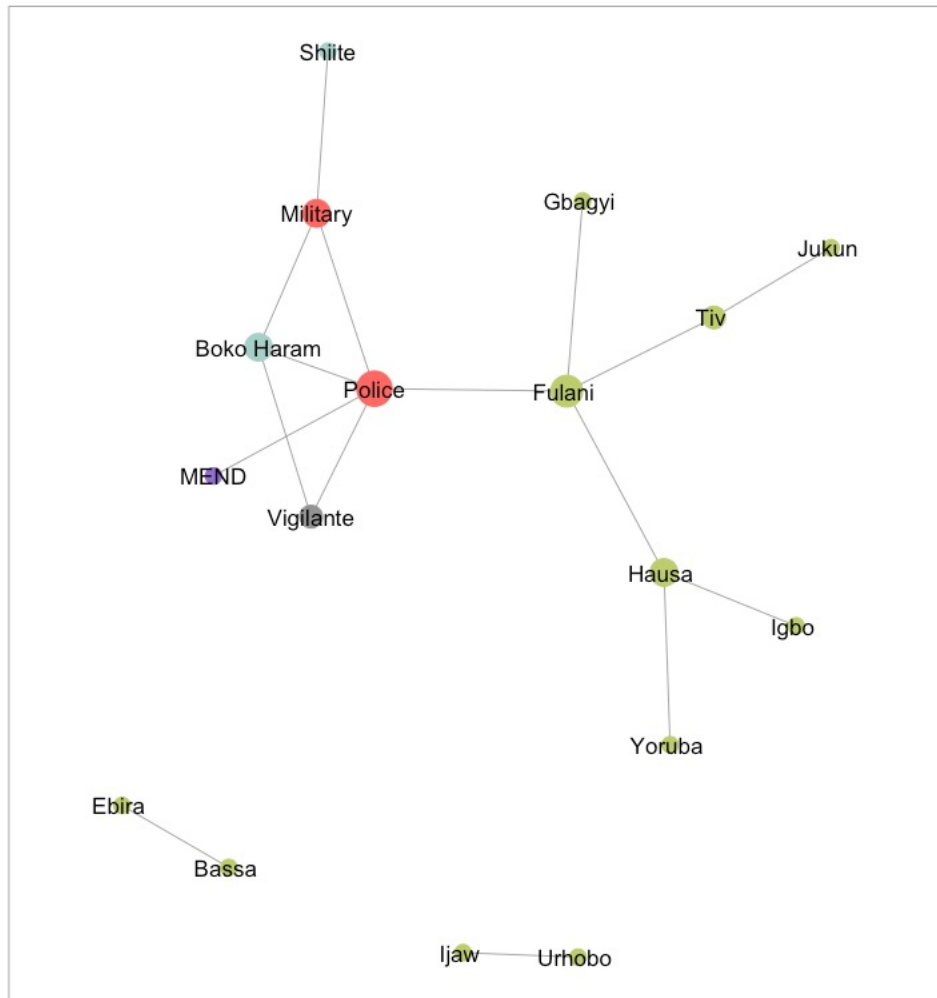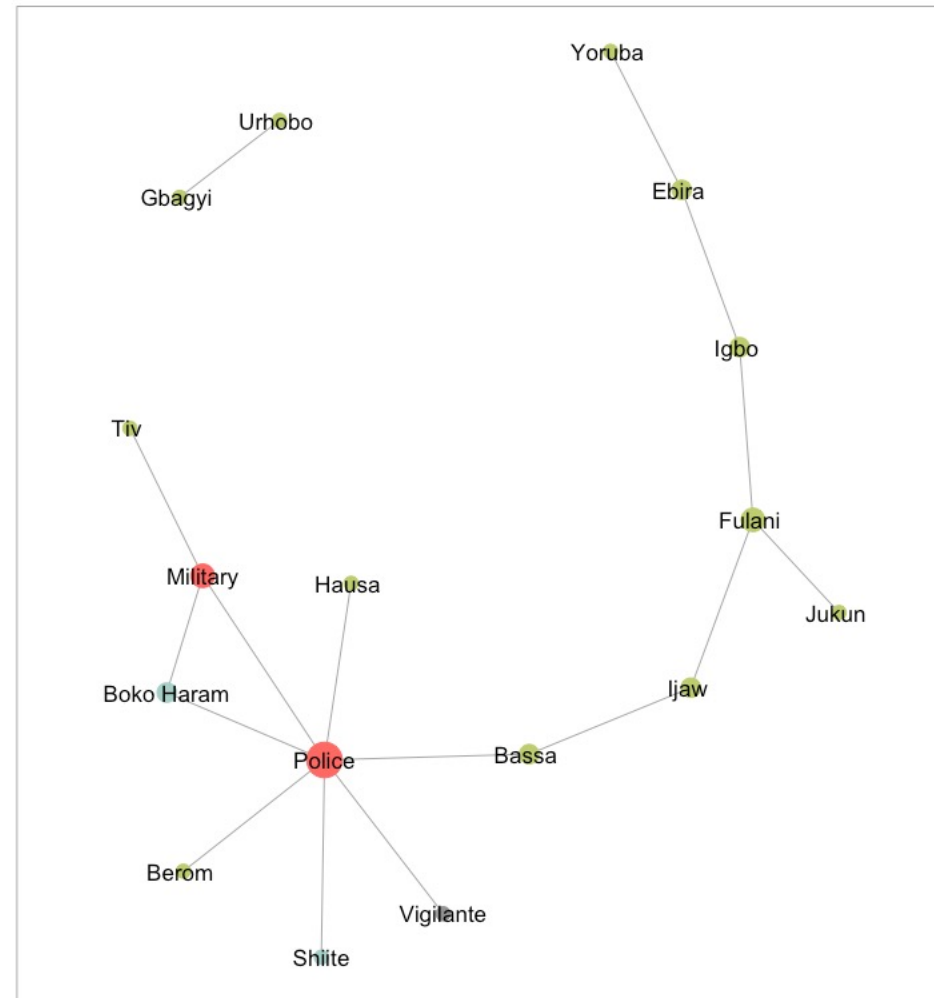
Simulated

Simulation 2016

Questions?

Code Repository

Replication files are available on Github under the following link:

https://github.com/mschalberger/conflict_network_nigeria.git

Refrences

Dorff, C., Gallop, M., & Minhas, S. (2020). Networks of Violence: Predicting Conflict in Nigeria. In The Journal of Politics (Vol. 82, Issue 2, pp. 476–493). University of Chicago Press. https://doi.org/10.1086/706459

Cranmer, S. J., Desmarais, B. A., & Morgan, J. W. (2020). Inferential Network Analysis (pp. 35–66). Cambridge University Press. https://doi.org/10.1017/9781316662915

Cranmer, S. J., & Desmarais, B. A. (2011). Inferential network analysis with exponential random graph models. *Political Analysis*, *19*(1), 66-86. https://doi.org/10.1093/pan/mpq037

Hunter, D. R., Krivitsky, P. N., & Schweinberger, M. (2012). Computational Statistical Methods for Social Network Models. In Journal of Computational and Graphical Statistics (Vol. 21, Issue 4, pp. 856–882). Informa UK Limited. https://doi.org/10.1080/10618600.2012.732921

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. In Social Networks (Vol. 29, Issue 2, pp. 173–191). Elsevier BV. https://doi.org/10.1016/j.socnet.2006.08.002