

# Modeling Social Dynamics in Directed Networks: An ERGM Analysis with Structural and Attribute-Based Predictors on the EIES Dataset

## Assignment 2

Agnese Cervino (agnese.cervino@studenti.unitn.it)

Damiano Orlandi (damiano.orlandi@studenti.unitn.it)

Maria Amalia Pelle (mariaamalia.pelle@studenti.unitn.it)

Maria Starodubtseva (mariia.starodubtseva@studenti.unitn.it)

## Introduction

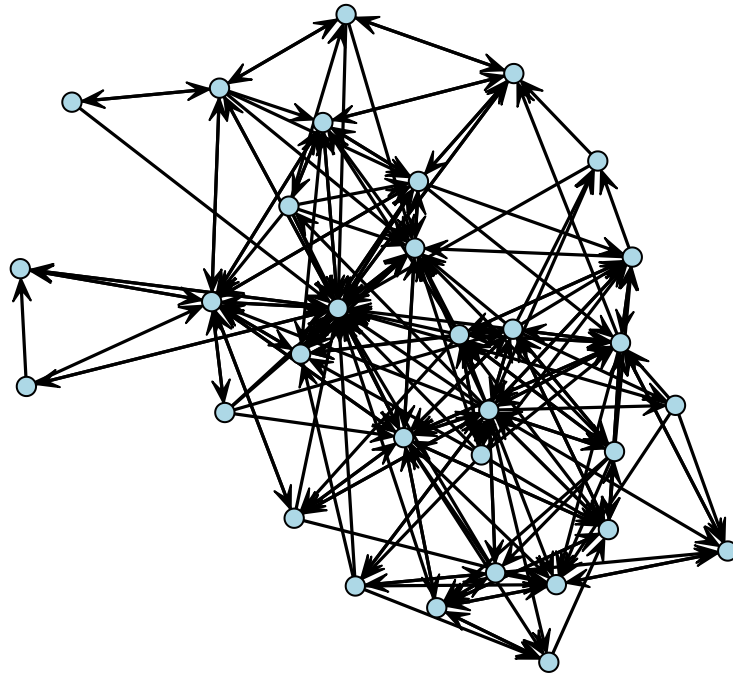
This assignment will explore the EIES dataset [1], which captures the dynamics of computer-mediated communication among a group of 32 network scientists. This dataset was part of an early experiment in the 1970s to study how a new communication medium influenced social relationships. The data consists of messages exchanged between academics interested in interdisciplinary research, collected at two distinct time periods. Specifically, we will focus on the acquaintance relationships at time 2, where individuals rated their connections on a scale from 0 (person unknown) to 4 (close personal friend).

The aim of examining this data is to understand the social relationships within this network and explore any potential patterns or trends that arise from these interactions, particularly in relation to the roles of gender and friendship within this academic context. To analyze these dynamics, we will employ an Exponential Random Graph Model (ERGM), leveraging both structural and attribute-based predictors to better understand the factors influencing the formation of ties in this directed network.

## Data Preparation

To analyze this network the values were dichotomized, transforming the network into a binary format where connections representing at least “friend” (3 or 4) are marked as 1, and all others (0, 1, or 2) are marked as 0. Additionally, later each individual is associated with a gender attribute, which will provide further insight into the network structure.

## Friendship network



Looking at the network it appears that some nodes display a significantly higher degree of connectivity than others.

### Part 1

Three models are fitted and run:

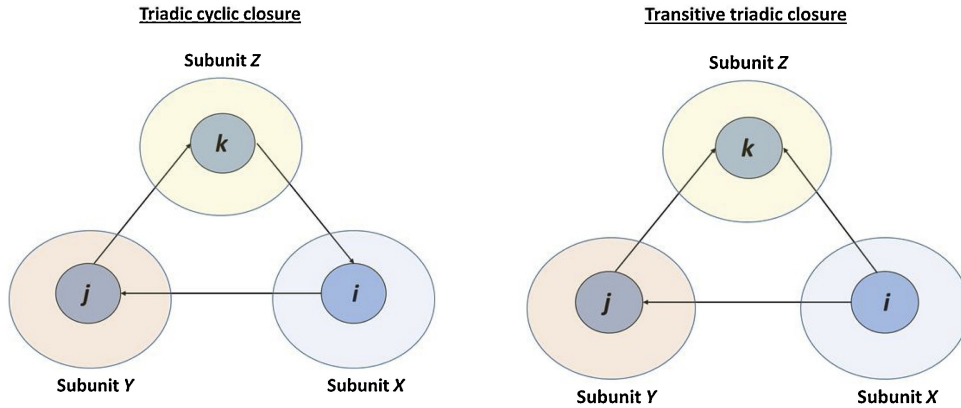
- Model 1: Purely structural model (no attributes)
- Model 2: Dyadic independent model with the attributes that were provided
- Model 3: A full model taking into consideration the structure and attributes

#### Model 1

This model takes into consideration only structural features of the network such as:

- Mutual connections (`mutual`): The presence of mutual connections, where both nodes are connected to each other.

- **Indegree centrality (`gwidegree`):** This measure captures the number of incoming connections (or edges) to a particular node within the network. It reflects the node's importance or influence, as a higher indegree centrality often indicates that the node is a significant recipient of connections, resources, or attention in the network. In our case, this metric identifies how many people consider a particular individual as their friend, showing the popularity and social status of that person.
- **Outdegree centrality (`gwodegree`):** Outdegree centrality, conversely, represents the number of outgoing connections from a node. This metric highlights the node's activity or outreach capacity within the network, showing how extensively it connects to other nodes. In our context, this metric indicates how many people a particular individual considers as friends, reflecting their social engagement and willingness to establish relationships.
- **Transitivity (`dgwesp(type="OTP")`):** Transitivity in directed networks measures the extent to which directed relationships between three nodes form closed loops. A triple of nodes  $(i, j, k)$  is considered transitive if there exists a directed edge from  $i$  to  $j$ , from  $j$  to  $k$ , and also from  $i$  to  $k$ . The transitivity of a directed network is often quantified using a clustering coefficient, defined as the fraction of directed two-paths  $(i \rightarrow j \rightarrow k)$  that are closed by a direct edge from  $i$  to  $k$ . This metric accounts for the directionality of connections, capturing the cohesive structure in directed networks [1].
- **Cyclicity (`dgwesp(type="ITP")`):** Cyclicity refers to the presence of cyclical triads or loops within the network. A cycle occurs when a sequence of connections forms a closed loop that returns to the starting node  $(i \rightarrow j \rightarrow k \rightarrow i)$ . Cyclicity provides valuable insights into feedback mechanisms, recurrent processes, or potential redundancies within the network's structure [1].



[2]

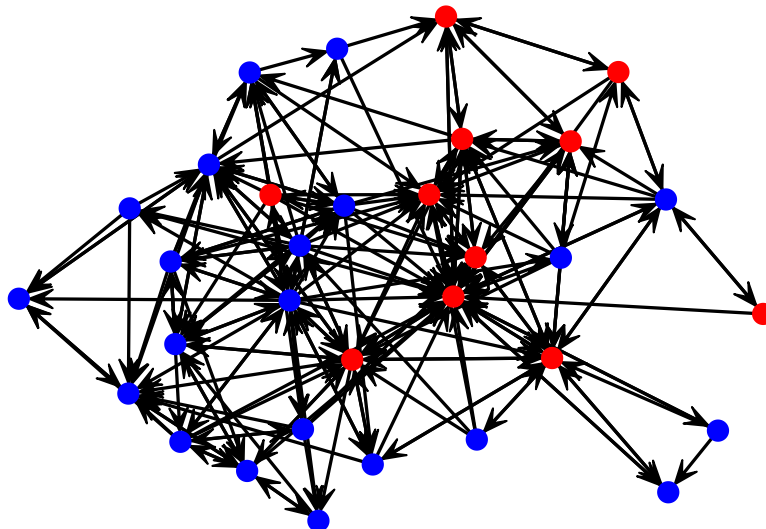
## Model 2

The following model takes into consideration mutual ties and a gender attribute assigned to each node, computing:

- Covariate for the receiver's gender (`nodeicov("Gender")`): A parameter that measures the effect of the receiver's gender on the probability of being targeted by interactions or connections. For example, it could account for whether one gender is more likely to be chosen as a friend by others in the network.
- Covariate for the sender's gender (`nodeocov("Gender")`): A parameter that captures how the sender's gender influences their likelihood of initiating interactions or connections. For example, it might model whether one gender is more active in sending connections (naming friends).
- Homophily (`nodematch("Gender")`): A measure of the tendency for individuals to form connections with others who share the same gender. This reflects the preference for same-gender interactions within the network.

The network with gender variable looks like this:

**Gender**  
■ Female  
■ Male



### Model 3

This model combines both structural effects (**Model 1**) and attribute effects (**Model 2**). It includes the following parameters: mutual connections, indegree and outdegree centrality, transitivity, cyclicity, covariates for the sender’s and receiver’s genders, and homophily.

It provides a fuller understanding view of the network dynamics by incorporating node attributes and structural configurations simultaneously.

## Part 2

In this section, we present the results from three models estimated using Exponential Random Graph Models (ERGMs), aiming to capture the underlying mechanisms that drive the formation of connections in the presented social network. We discuss the estimated coefficients, their significance, and the implications for understanding the network structure.

Particularly we describe **Model 3**, which incorporates structural and attribute effects. **Model 1** and **Model 2** are also reviewed to highlight any differences that emerge as significant or insignificant in **Model 3**.

### Results Table

Term	Model 1			Model 2			Model 3		
	Est	SE	Sign	Est	SE	Sign	Est	SE	Sign
Edges	-4.8232	0.3524	***	-2.8028	0.1755	***	-5.0750	0.3464	***
Mutual	2.5927	0.3602	***	2.6742	0.2915	***	2.5598	0.3590	***
Gwidegree	2.0618	1.2289	.	-	-	-	2.4091	1.3509	.
Gwodegree	4.4934	2.731	.	-	-	-	3.4516	2.3692	.
Dgwesp-OTP	1.9615	0.2156	***	-	-	-	1.8315	0.2168	***
Dgwesp-ITP	-0.6201	0.1088	***	-	-	-	-0.5139	0.1167	***
Nodeicov	-	-	-	1.1593	0.2034	***	0.5644	0.1622	**
Nodeocov	-	-	-	-0.4204	0.2115	*	-0.2293	0.1840	.
Nodematch	-	-	-	0.5238	0.1434	***	0.3199	0.1014	**

*Note:* ‘-’ indicates not estimated

*Significance Levels:* ‘\*\*\*’:  $p < 0.001$ ; ‘\*\*’:  $p < 0.01$ ; ‘\*’:  $p < 0.05$ ; ‘.’:  $p < 0.1$

First, the **edges** term is statistically significant across all models, indicating that the network’s density is not random and that the observed number of connections significantly deviates from what would be expected in a random graph.

Next, we observe that the **mutual** predictor is statistically significant across all models, indicating that the presence of mutual connections is essential in explaining the network structure. The positive coefficient (2.5598,  $p < 0.001$ ) for **mutual** in **Model 3** demonstrates

a tendency toward reciprocity in friendships. This suggests that individuals are more likely to form connections with others who consider them friends, aligning with the idea that friendship is a mutual relationship.

The positive coefficients for `gwidegree` (2.4091,  $p < 0.1$ ) and `gwodegree` (3.4516,  $p > 0.1$ ) in **Model 3** suggest that individuals with higher indegree and outdegree centrality are more likely to form connections. This indicates that individuals who are considered friends by others (indegree) and who consider others as friends (outdegree) tend to form more connections. Moreover, the magnitude of the coefficient for `gwodegree` is higher than that of `gwidegree`, suggesting that outdegree centrality would have a stronger effect on the formation of connections than indegree centrality. So individuals who are naming more friends could play a relatively more significant role in the network formation process. Estimated coefficients are not statistically significant, though.

The parameter `dgwesp(type = "OTP")` measures transitivity by looking at how nodes with outgoing ties contribute to closed triadic structures. The positive coefficient (1.8315,  $p < 0.001$ ) in **Model 3** indicates that presence of transitive triadic structures is positively associated with the likelihood of forming connections in this network. In a friendship network, transitive triads foster connected groups of friends, where friends of friends are more likely to become friends themselves.

On the contrary, the negative coefficient for `dgwesp(type = "ITP")` (-0.5139,  $p < 0.001$ ) in **Model 3** suggests that the presence of cyclic triads in the network has a negative effect on the formation of connections. This implies that individuals who are part of cyclic triads are less likely to form connections with others. In a friendship network, people in cyclic triads are less likely to create new connections with others because the cyclic nature might reduce the opportunities or motivations for forming additional friendships.

The `nodeicov` term in **Model 3** has a positive coefficient (0.5644,  $p < 0.01$ ), indicating that female scientists (coded as 1) are more likely to be named as friends by others compared to male scientists (coded as 0). The result suggests that females are perceived as more attractive targets for friendship nominations in the network, potentially reflecting social norms or behaviors, where females might have more influence and are more likely to be named as friends. They might be perceived as more approachable or friendly.

The `nodeocov` term in **Model 2** has a significant negative coefficient (-0.4204,  $p < 0.05$ ) which becomes -0.2293 and insignificant in **Model 3**. This indicates that, initially, females appeared less likely to name others as friends compared to males. However, after accounting also for structural effects in **Model 3**, this relationship weakens and is no longer statistically significant. This suggests that the apparent difference in females' likelihood of naming friends in **Model 2** may have been confounded by other network dynamics, such as degree distributions.

Finally, the `nodematch` term in **Model 3** has a positive coefficient (0.3199,  $p < 0.001$ ), indicating that nodes with the same gender are more likely to form connections in the network. This reflects the tendency for homophily in friendship networks, where females are more likely to be friends with females and males with males. This could be due to shared experiences, interests, or social norms that lead to the formation of same-gender friendships.

It is important to note the gender imbalance (21 males, 11 females) in the network. This imbalance may lead to less precise estimates for female-specific connections due to the smaller size of the female group.

### Part 3

Evaluating a model's goodness of fit is crucial for validating its alignment with observed data and ensuring reliable inference. For Model 3, diagnostics such as histograms, trace plots, and boxplots of simulated values offer valuable visual insights into model performance and convergence, confirming well-behaved posterior distributions and robust parameter estimates.

To assess the model's goodness of fit, we conduct a Goodness-of-Fit (GOF) analysis using 2000 simulations. The table below summarizes the model fit for specific terms (e.g., edges, mutual ties, covariates) by comparing observed values with those from the simulated networks:

##	obs	min	mean	max	MC p-value
## edges	204.00000	152.00000	204.54400	260.00000	0.999
## mutual	60.00000	38.00000	60.71050	87.00000	0.993
## gwideg.fixed.0.3	41.74622	37.74549	41.74012	43.13550	0.844
## gwodeg.fixed.0.3	42.57393	39.35834	42.58105	43.17246	0.759
## gwesp.OTP.fixed.0.5	271.15878	170.94551	272.64087	378.56748	0.966
## gwesp.ITP.fixed.0.5	217.35306	118.01650	219.53260	331.32618	0.964
## nodeicov.Gender	99.00000	50.00000	98.72800	139.00000	1.000
## nodeocov.Gender	72.00000	42.00000	72.20300	105.00000	1.000
## nodematch.Gender	125.00000	80.00000	125.98300	206.00000	1.000

Each row represents a model statistic, displaying the observed value, the range of simulated values (minimum and maximum), the mean of the simulations, and the Monte Carlo (MC) p-value.

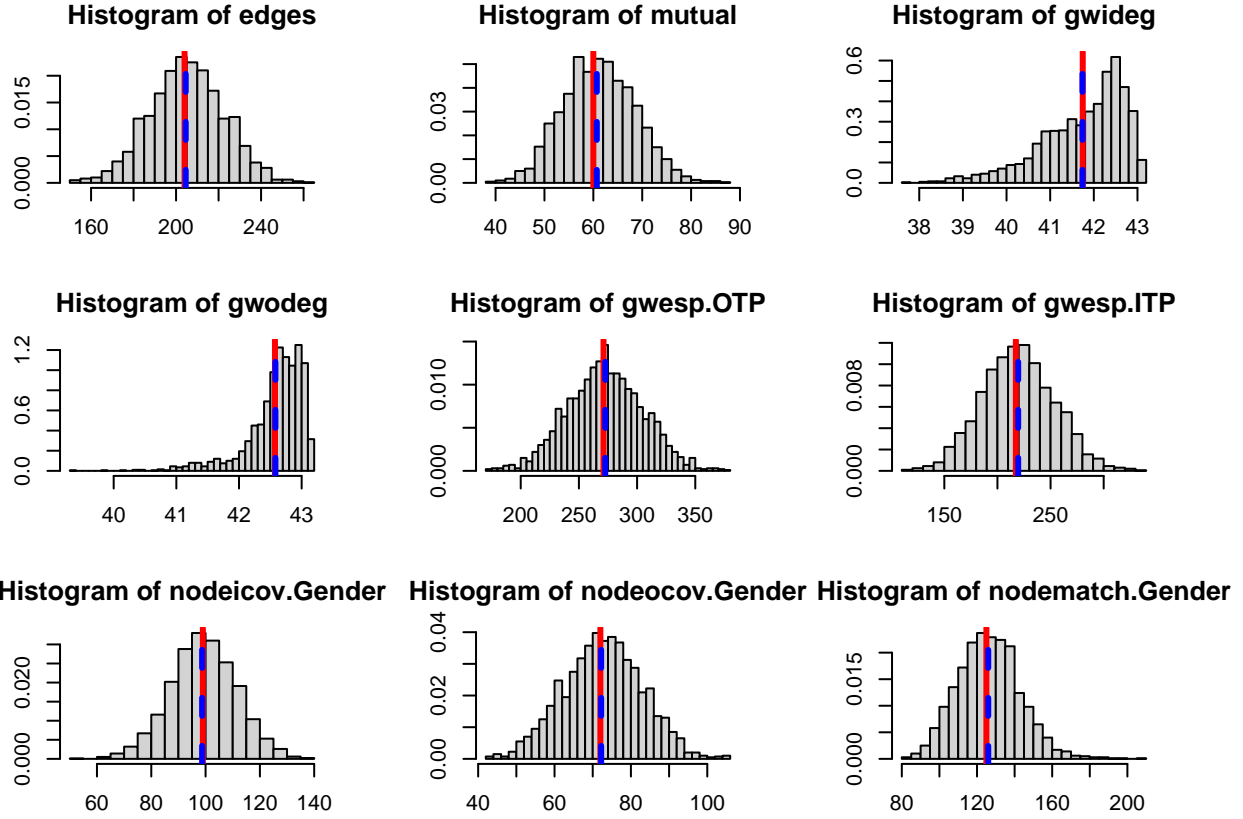
The observed values align closely with the simulation means, indicating an excellent overall fit. The high p-values (all close to 1) confirm that the observed statistics fall well within the range of expected values under the model.

The lowest p-value corresponds to `gwodeg.fixed.0.3` (outdegree centrality). However, the difference between the observed value and the simulation mean is only 0.01. This is likely due to the narrow range of simulated values (39.36 to 43.17), suggesting low variability in this statistic.

Conversely, while `gwesp.ITP.fixed.0.5` deviates the most from its simulation mean, with a difference of 2.17 points, its p-value remains high at 0.964. This is because the simulated values span a wider range (118.02 to 331.33), including the observed value within a broader interval.

## Histograms

A histogram for each parameter is plotted to display the distribution of simulated values. Two reference lines are included: the observed value in red and the simulation mean in blue.



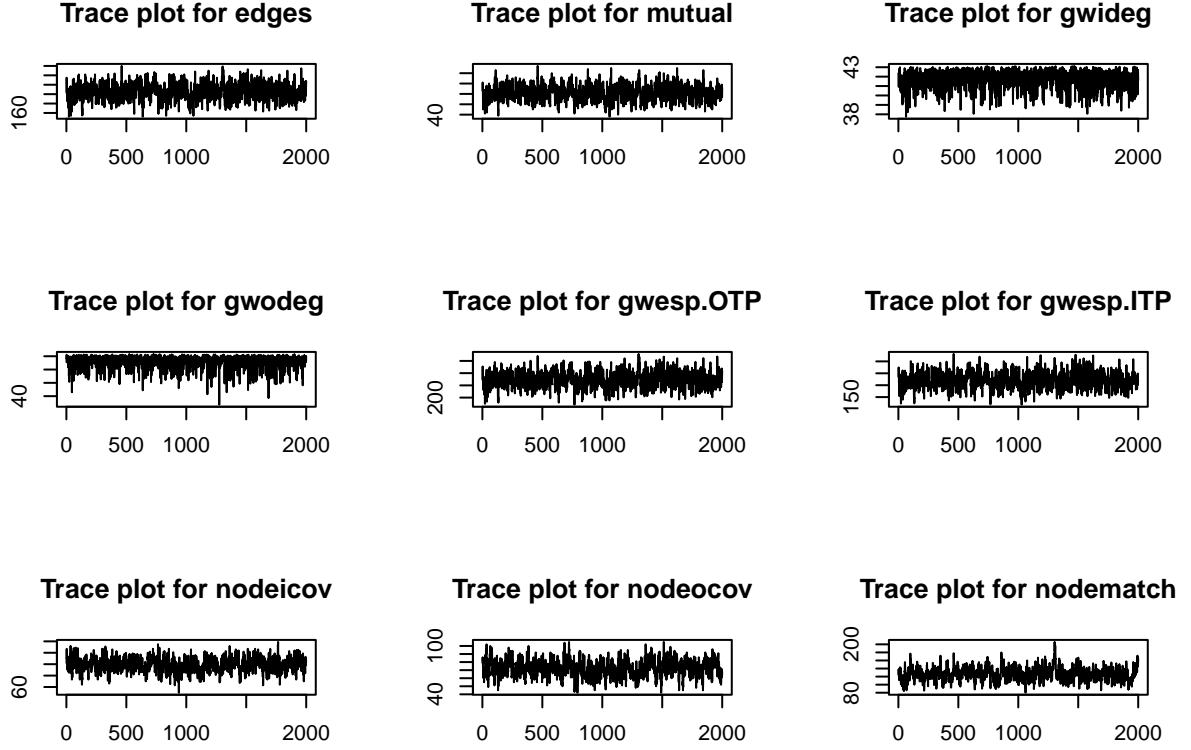
Each parameter exhibits its own unique distribution. Most simulated values follow a normal distribution, while the ones deviating the most are outdegree and indegree centrality. Outdegree and indegree centrality usually have right-skewed distributions, where a few nodes are highly connected while most have low centrality. This deviation from normality reflects the network's structural constraints and the presence of hubs. Centrality measures are also sensitive to network sparsity, further contributing to this pattern.

Visual inspection shows again that the simulated values (blue dashed line) are consistent with the observed values (red line) for each parameter, further confirming the model's goodness of fit.

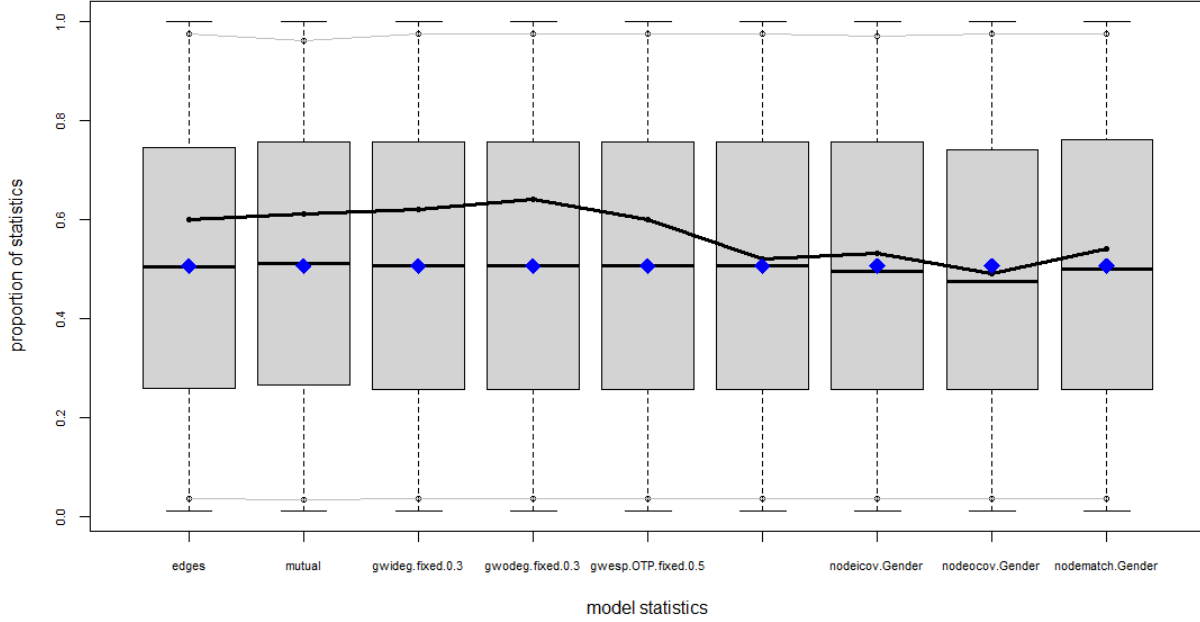
## Trace Plots

We plot the trace plots to visualize the behavior of simulated parameter values across iterations in the model. The Y-axis represents the parameter values, while the X-axis corresponds to the iteration number.





Each plot shows the parameter estimates fluctuating within a consistent range of values and around a stable mean, indicating good stationarity and convergence of the MCMC chains. The absence of strong trends or drifts across iterations suggests that the chains have explored the parameter space adequately. Overall, the trace plots provide strong evidence of proper MCMC behavior, supporting the reliability of the model's parameter estimates and the validity of the model fit for the observed network data. We can notice taht the scale of the Y-axis varies across parameters, reflecting differences in their estimated magnitudes. For example, `gwideg` and `gwodeg` plots exhibit narrower fluctuations, reflecting tightly constrained degree distributions in the network.

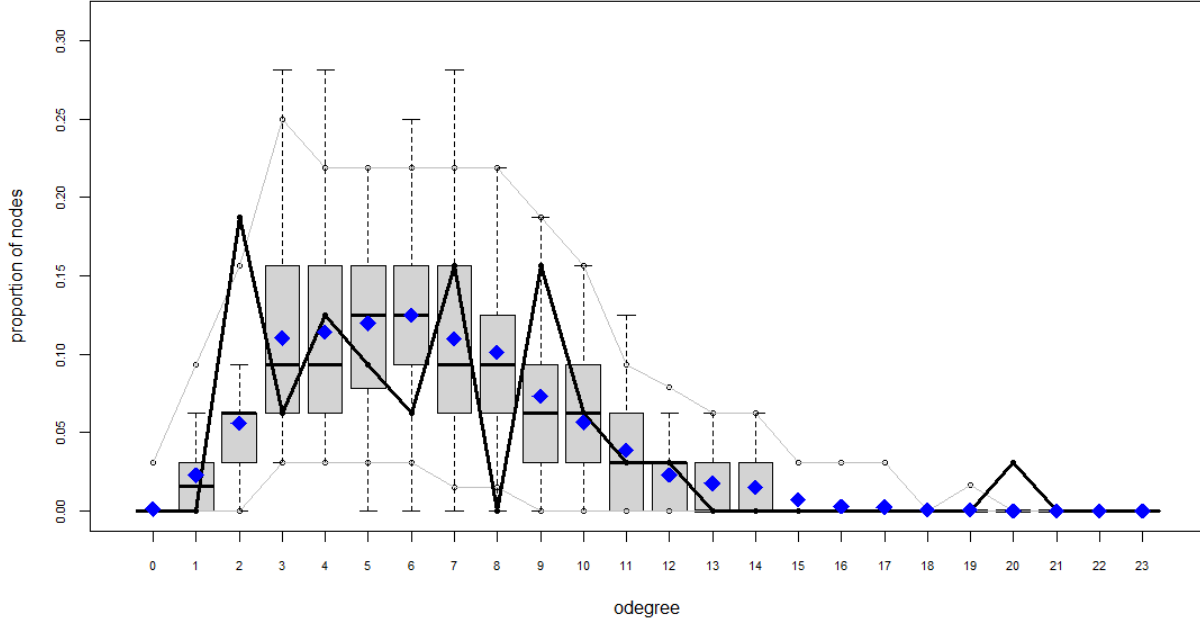


The plot assesses how well the simulated networks replicate key statistics observed in the real network. The X-axis represents the specific structural statistics used to evaluate the model’s fit, such as edges, mutual ties, weighted degree distributions (`gwideg.fixed.0.3`, `gwodeg.fixed.0.3`), transitivity (`gwesp.OTP.fixed.0.5`), and attribute-based patterns like gender influence and homophily, while the Y-axis shows the normalized proportion of each statistic.

The model shows a slight tendency to underestimate density, reciprocity, connectivity (both indegree and outdegree) and transitivity, as evidenced by the lack of overlap between the black line (observed values) and the mean for the `edges`, `mutual`, `gwideg.fixed.0.3`, `gwodeg.fixed.0.3` and `gwesp.OTP.fixed.0.5` statistics. The slight underestimation highlights that these features might be shaped by latent factors, such as shared norms or goals, proximity, or collaborative culture, that the model fails to fully incorporate.

In contrast, cyclicity, gender influence, and homophily are well-fitted by the model, as indicated by an alignment of the simulated means and the real network values. Indeed, these factors are more straightforward to model, as they follow consistent patterns in social relationships.

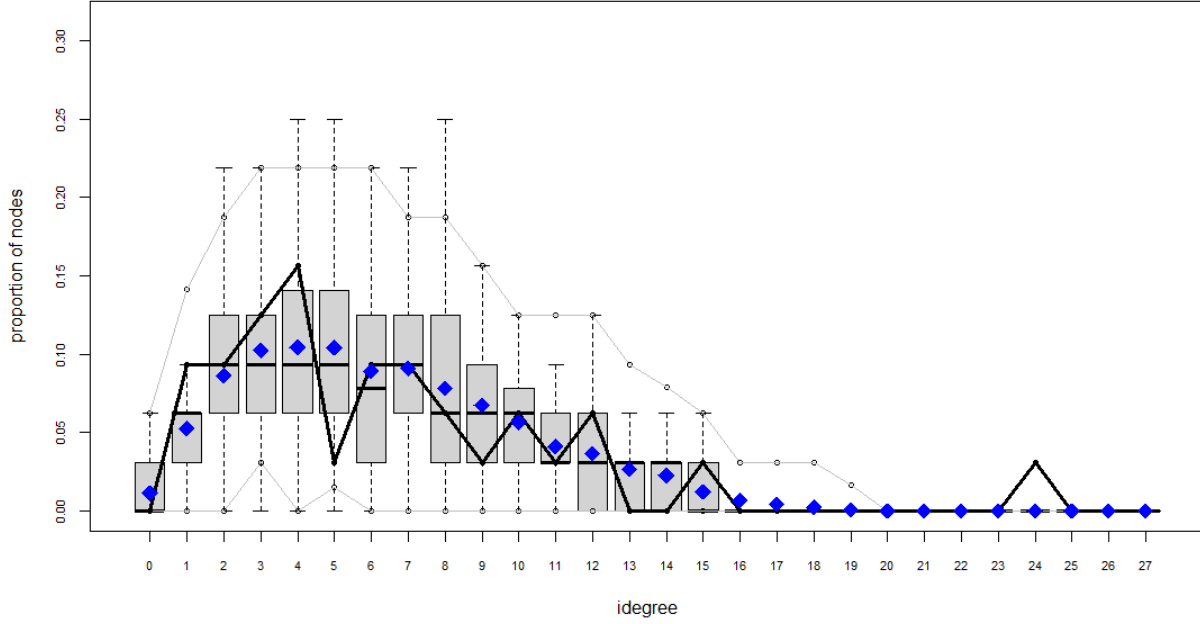
Overall, the analysis suggests that the real network is well-represented by the model, identifying both the structural features and the individual-level attributes.



The plot represents the simulated and observed networks' outdegree distributions, with the X-axis visualizing the number of outgoing connections per node, and the Y-axis illustrating the proportion of nodes with each outdegree value.

We can see clearly discrepancies between the networks at specific outdegree values (particularly at 2, 3, 5, 6, 7, 8, 9 and 20) indicating the model's tendency to generalize the heterogeneity of the real network's outdegree distribution. In the real network, outdegree values of 2, 9, and 20 are much higher than those in the model. On the other hand, outdegree values of 3, 5, 6, and 8 are overestimated in the simulations on average.

The distributions of simulated outdegree values are generally consistent with the real network, but on average across 2000 simulations the model tends to underestimate or overestimate some outdegree values of the real network.



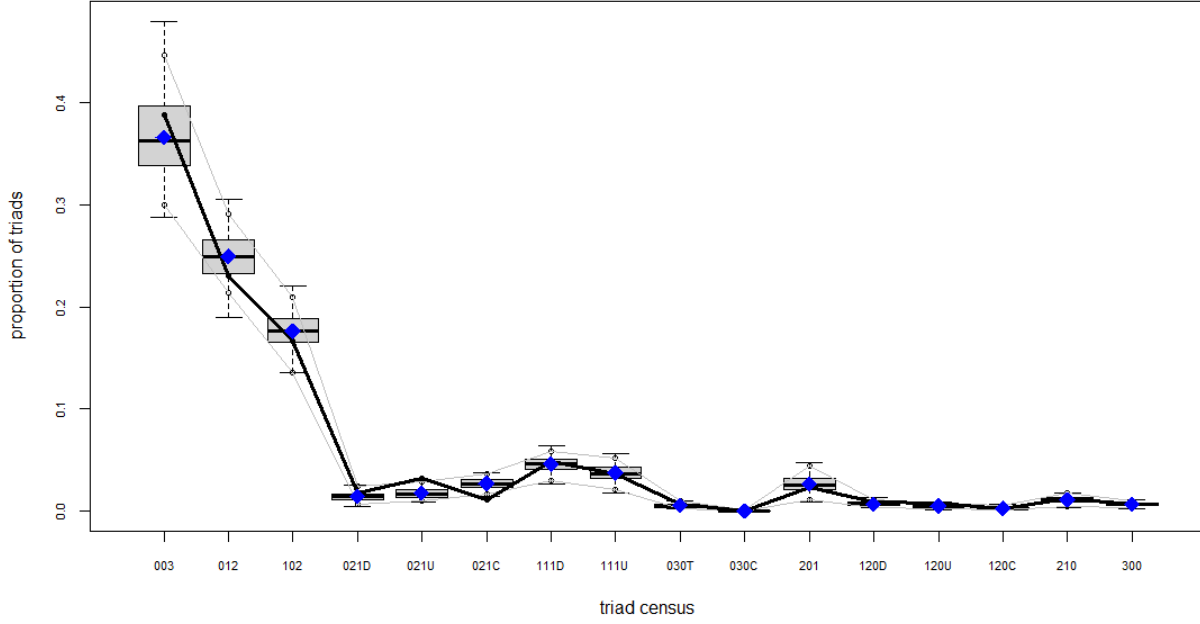
This plot illustrates the simulated and observed networks' indegree distributions by representing the number of incoming connections per node, and the proportion of nodes with each indegree value.

In this case, the model generally performs well in replicating the real-world network, with fewer discrepancies than seen in the outdegree distribution. However, some underestimations and overestimations still occur at specific indegree values.

Specifically, at indegree 1 and 4 the real network has a much higher proportion of nodes with incoming connections than the model does on average. Similarly, at indegree 3, 12 and 15, the real network shows slightly higher proportions (not exceeding the borders of the boxplots), suggesting that moderate levels of incoming ties are slightly more prevalent in reality. Interestingly, at indegree 24, the real network shows a peak (while simulated values are around 0), pointing to highly centralized nodes that likely act as hubs for incoming connections.

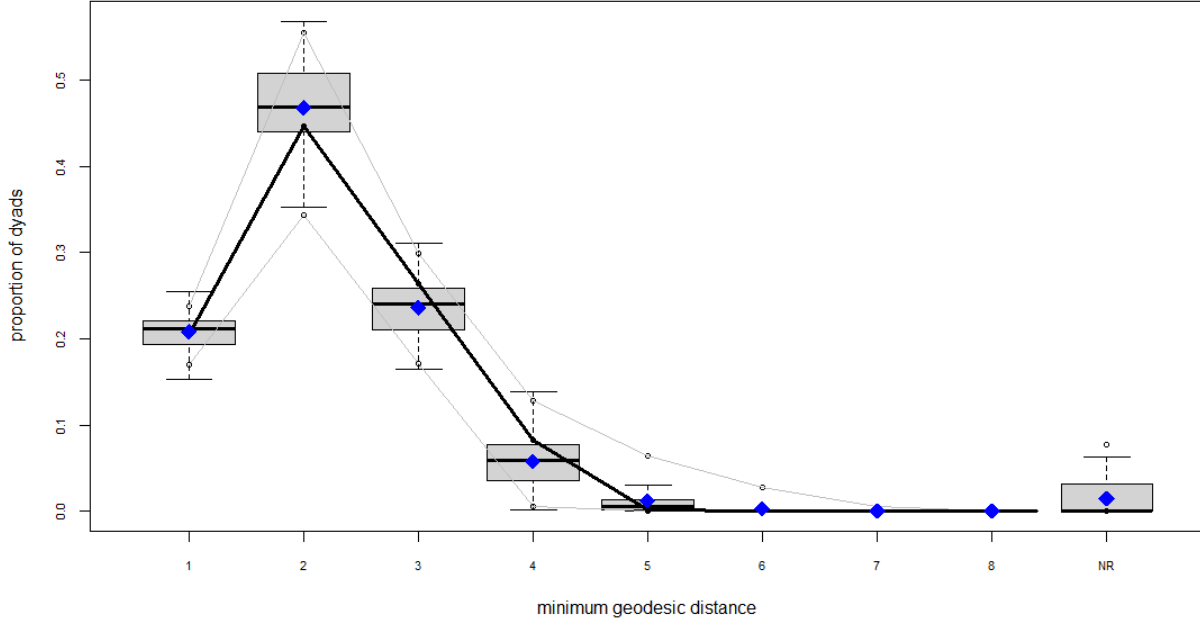
In contrast, at indegree 5, 9, 11, 13 and 14 the model on average estimates higher indegree values than observed in the real network, overestimating the prevalence of nodes with these values.

Overall, the model captures the general shape of the indegree distribution, but some discrepancies exist at specific indegree values, indicating that the model may not fully capture the heterogeneity of incoming connections in the real network.



The plot compares the proportions of different triad types in the real and simulated networks. The X-axis represents the triad census while the Y-axis shows the proportion of each triad type relative to the total number of triads in the network. It is straightforward that the model is almost perfectly representing the real world network, specifically simple triads (003, 012, 102) and low-frequency complex triads (120D, 120C etc.). However, the only exceptions are given by the triads 021U and 021C, which are respectively underestimated and overestimated. Overall, the plot respects the usual distribution trends, capturing the dominance of simpler configurations and the decline in frequency as triads become more complex, which suggests the model is effectively reproducing the structural tendencies of the real network.

### Goodness-of-fit diagnostics



The plot illustrates networks' minimum geodesic distance distributions, where the X-axis represents the geodesic distance, and the Y-axis displays the proportion of dyads at each distance. In this case, the model performs well in replicating the real-world network, particularly for longer distances. Overall, the simulation performs well in representing the network's sparsity.

### Triadic census

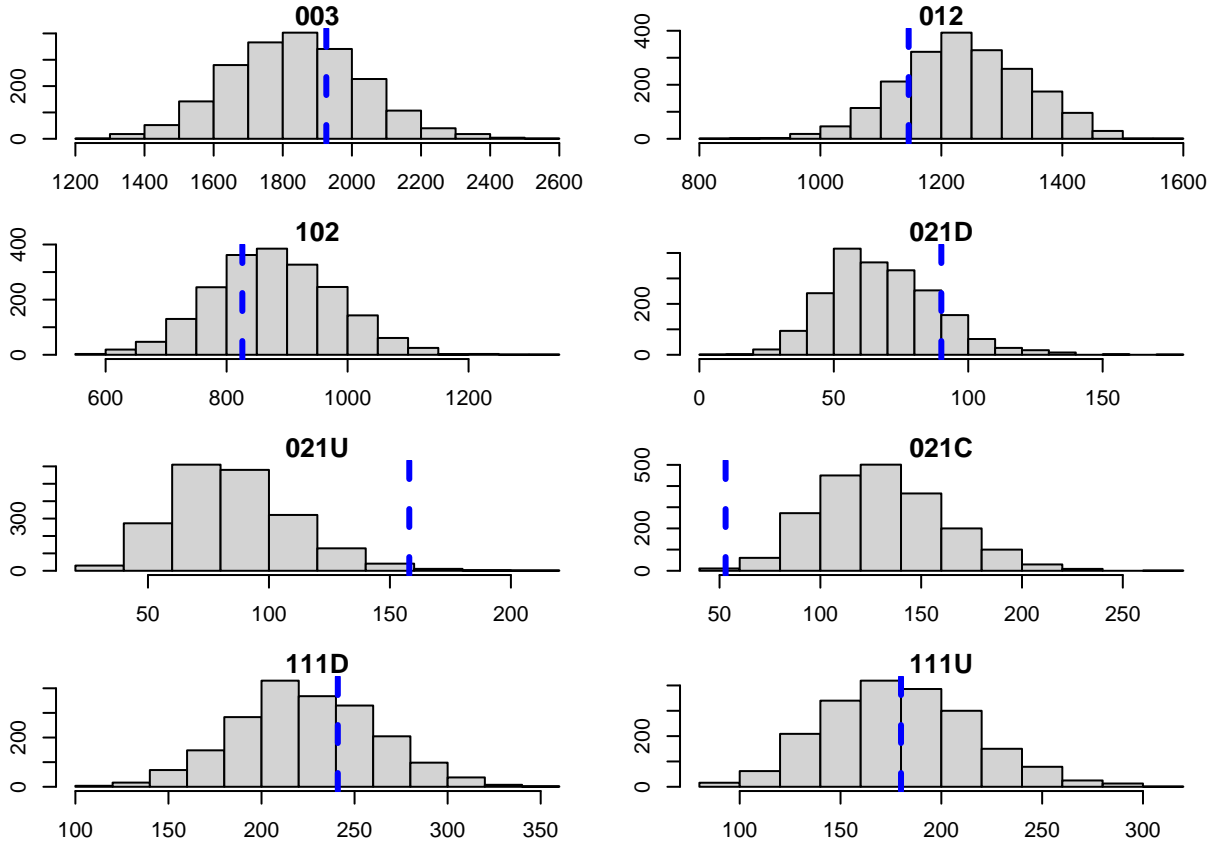
In a directed network there are sixteen possible triads. In this part we count the number of each type of triad present in the network. The triads are labelled  $abcZ$  where  $a$  is the number of reciprocated ties,  $b$  is the number of unreciprocated ties and  $c$  is the number of null ties. The  $Z$  term is a letter (U,C,D or T) used to differentiate between different triads in which these numbers are the same. For example, **003** represents the empty triad, **012** represents a triad with a single directed edge and **102** represents a triad with a single mutual edge.

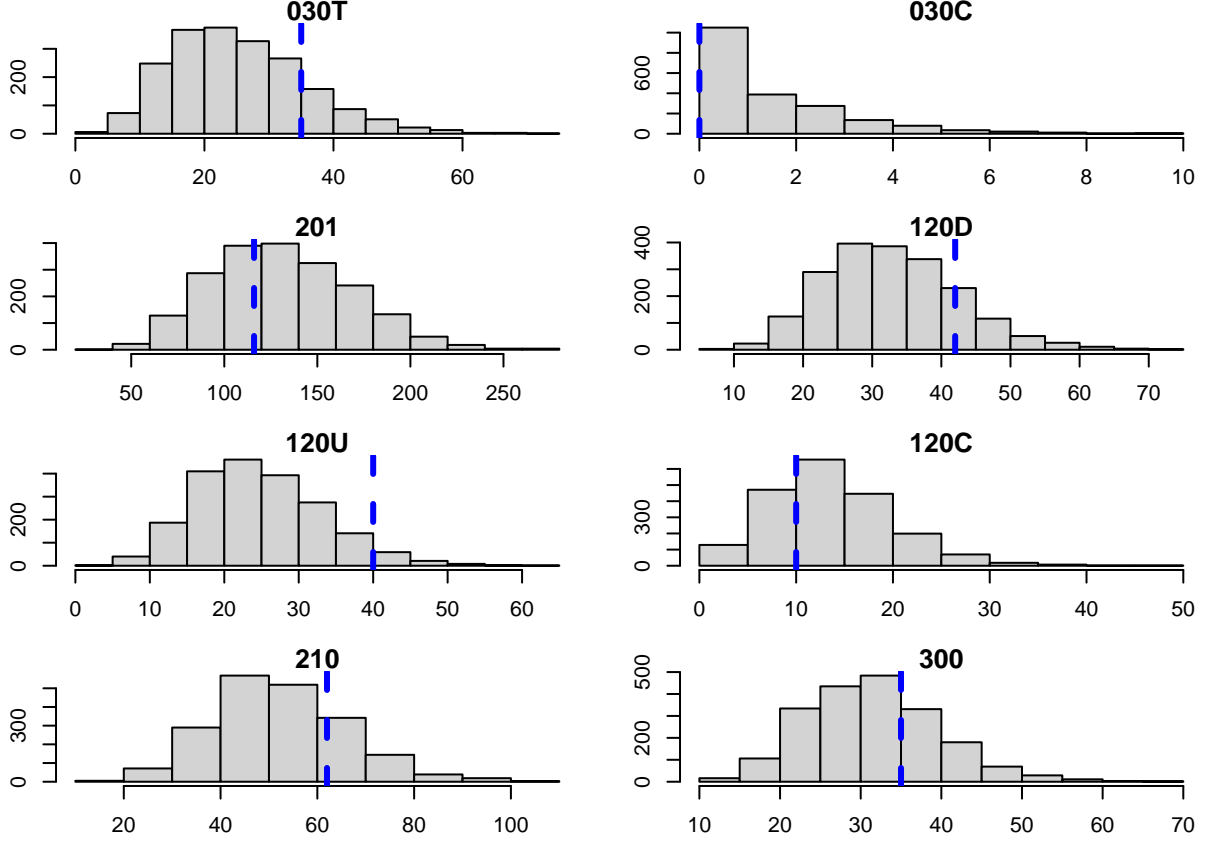
- 003:  $A, B, C$ , empty triad
- 012:  $A \rightarrow B, C$
- 102:  $A \leftrightarrow B, C$
- 021D:  $A \leftarrow B \rightarrow C$
- 021U:  $A \rightarrow B \leftarrow C$
- 021C:  $A \rightarrow B \rightarrow C$
- 111D:  $A \leftrightarrow B \leftarrow C$
- 111U:  $A \leftrightarrow B \rightarrow C$
- 030T:  $A \rightarrow B \leftarrow C, A \rightarrow C$

- 030C:  $A \leftarrow B \leftarrow C, A \rightarrow C$
- 201:  $A \leftrightarrow B \leftrightarrow C$
- 120D:  $A \leftarrow B \rightarrow C, A \leftrightarrow C$
- 120U:  $A \rightarrow B \leftarrow C, A \leftrightarrow C$
- 120C:  $A \rightarrow B \rightarrow C, A \leftrightarrow C$
- 210:  $A \rightarrow B \leftrightarrow C, A \leftrightarrow C$
- 300:  $A \leftrightarrow B \leftrightarrow C, A \leftrightarrow C$ , completely connected

The histograms illustrate the distribution of specific triad types in simulated datasets, with the X-axis showing the frequency of each triad type and the Y-axis representing the number of simulation runs in which those counts occurred.

Each bar indicates how often a particular triad count appeared in the simulations. The blue dashed line marks the observed count of the respective triad type in the actual network data. By comparing the position of the blue line to the histogram, we can assess how well the simulated data aligns with the observed data, providing insights into the model's goodness of fit.





We can observe that the model generally performs well in replicating the real-world network. The observed values are close to the peak of the histograms for most triad types, indicating that the model captures the network’s structural characteristics effectively. However, there are some discrepancies for specific triad types, such as 021D, 021U, 021C, 030T and 120U.

The plot for 030C, which represents unidirectional cycles within the network, shows an observed value that is extremely low. This indicates that such configurations are unlikely in the real network, reflecting a strong tendency toward mutuality in friendship relationships.

For 021D triad ( $A \leftarrow B \rightarrow C$ ) the simulated values are mostly lower than the observed value, suggesting that the model underestimates the prevalence of this triad type in the real network. There are more triads of friendships with two outgoing ties from the central node than the model predicts. Similarly, for 021U triad ( $A \rightarrow B \leftarrow C$ ) the simulated values are mostly lower than the observed value, underestimating the prevalence of this triad type in the real network. There are more triads of friendships with two incoming ties to the central node than the model predicts. It indicates potential structural features in the friendship network, such as people being more central or influential in forming relationships than the model accounts for.

On the contrary, for 021C triad ( $A \rightarrow B \rightarrow C$ ) the simulated values are mostly higher than the observed value, overestimating the prevalence of this triad type in the real network. This suggests the model predicts more fully transitive relationships than are actually observed in



the network. In reality, such direct, chain-like friendships are less common than the model assumes.

For the 030T triad ( $A \rightarrow B \leftarrow C, A \rightarrow C$ ), which represents transitive triadic closure, the lower simulated values compared to the observed suggest that the model underestimates the prevalence of this triad type in the real network. This indicates that there may be more transitive friendship triads in the real network than the model predicts.

Finally, for 120U triad ( $A \rightarrow B \leftarrow C, A \leftrightarrow C$ ) the simulated values are mostly lower than the observed value, underestimating the prevalence of this triad type, meaning it predicts fewer instances of such mixed-directional connections than are actually observed. In real friendship networks, people often form connections where both incoming and outgoing ties are present, reflecting a more balanced or reciprocal structure between nodes.

These results suggest a network that favors mutuality while avoiding cycles and non-reciprocal relationships.

## Conclusions

This study analyzed the friendship network of 32 scientists, revealing key factors that influence the creation of social ties within this academic community. With the help of an Exponential Random Graph Model (ERGM) and MCMC (Monte Carlo Markov Simulation), we simulated networks to replicate the EIES network structure and assess the model's goodness of fit. The models revealed significant tendencies for reciprocity, transitivity, and gender homophily, while also highlighting the influence of gender in both the formation and reception of ties. Although the models performed well in capturing core structural tendencies and individual-level attributes, some limitations emerged in replicating complex triadic structures and distributional variability.

## References

1. Freeman, S C and L C Freeman (1979). The networkers network: A study of the impact of a new communications medium on sociometric structure. Social Science Research Reports No 46. Irvine CA, University of California.
2. Tasselli, S., & Caimo, A. (2019). Does it take three to dance the Tango? Organizational design, triadic structures and boundary spanning across subunits. Social Networks, 59, 10-22.

## Appendix

```
library(readxl)
library(igraph)
library(sna)
library(network)
library(ergm)
library(knitr)
library(kableExtra)

setwd("/Users/maria/Desktop/ASNA")
raw_data <- read.csv("EIES_T2.csv", row.names = 1)
data <- ifelse(raw_data > 2, 1, 0)
net<- as.network(data, directed = TRUE)
par(mfrow=c(1,1))
par(mar=c(0,0,2,0))
gplot(net,
      vertex.cex = .8,
      displaylabels = FALSE,
      edge.col = "black",
      vertex.col = "lightblue",
      main = "Friendship network")

knitr::download_image(
  "https://ars.els-cdn.com/content/image/1-s2.0-S0378873318300169-gr1_lrg.jpg")

# Part 1
MOD1<-ergm(net ~ edges+mutual
            +gwidegree(decay=.3, fixed=T)
            +gwodegree(decay=.3, fixed=T)
            +dgwesp(type="OTP", decay=.5, fixed=T)
            +dgwesp(type="ITP", decay=.5, fixed=T),
            control=control.ergm(seed=101,
                                  MCMC.runtime.traceplot=FALSE,
                                  # MCMC.interval=64,
                                  # MCMC.samplesize=5000
            ),
            verbose=T)

gender_data <- read.csv("Data_Gender_12.csv",row.names = 1, sep = ";")
net%v%"Gender"<-gender_data
par(mar = c(0, 0, 1, 0))
node_colors <- ifelse(gender_data == 1, "red", "blue")
gplot(net,
```

```

vertex.cex = .8,
displaylabels = FALSE,
vertex.col = node_colors,
vertex.border = node_colors)

legend("topleft",
  legend = c("Female", "Male"),
  fill = c("red", "blue"),
  title = "Gender",
  title.font = 2,
  bg = "transparent", cex = 0.9, box.col = "transparent")

MOD2 <- ergm(net ~ edges + mutual +
  nodeicov("Gender") +
  nodeocov("Gender") +
  nodematch("Gender"),
  control = control.ergm(seed = 5,
    MCMC.runtime.traceplot = FALSE),
  verbose = TRUE)

MOD3<-ergm(net ~ edges+mutual
  +gwidegree(decay=.3, fixed=TRUE)
  +gwodegree(decay=.3, fixed=TRUE)
  +dgwesp(type="OTP", decay=.5, fixed=TRUE)
  +dgwesp(type="ITP", decay=.5, fixed=TRUE)
  +nodeicov("Gender")+nodeocov("Gender")
  +nodematch("Gender"),
  control=control.ergm(seed=102,
    MCMC.runtime.traceplot=FALSE),
  verbose=TRUE)

# Part 2
data <- data.frame(
  Term = c("Edges", "Mutual", "Gwidegree",
    "Gwodegree", "Dgwesp-OTP", "Dgwesp-ITP",
    "Nodeicov", "Nodeocov", "Nodematch"),
  Model1_Est =
    c(-4.8232, 2.5927, 2.0618, 4.4934, 1.9615, -0.6201, "-", "-", "-"),
  Model1_SE =
    c(0.3524, 0.3602, 1.2289, 2.7310, 0.2156, 0.1088, "-", "-", "-"),
  Model1_Sig =
    c("****", "****", ".", ".", "****", "****", "-", "-", "-"),
  Model2_Est =

```

```

      c(-2.8028, 2.6742, "-", "-", "-", "-", 1.1593, -0.4204, 0.5238),
Model2_SE =
      c(0.1755, 0.2915, "-", "-", "-", "-", 0.2034, 0.2115, 0.1434),
Model2_Sig =
      c("***", "***", "-", "-", "-", "-", "***", "*", "***"),
Model3_Est =
      c(-5.0750, 2.5598, 2.4091, 3.4516, 1.8315,
        -0.5139, 0.5644, -0.2293, 0.3199),
Model3_SE =
      c(0.3464, 0.3590, 1.3509, 2.3692, 0.2168,
        0.1167, 0.1622, 0.1840, 0.1014),
Model3_Sig =
      c("***", "***", ".", " ", "***", "***", "**", " ", "**")
)
kable(
  data,
  col.names = c("Term",
                 "Est", "SE", "Sign",
                 "Est", "SE", "Sign",
                 "Est", "SE", "Sign"),
  align = 'c',
  escape = FALSE
) %>%
  add_header_above(c(" " = 1,
                     "Model 1" = 3,
                     "Model 2" = 3,
                     "Model 3" = 3)) %>%
  kable_styling(full_width = FALSE, position = "center") %>%
  footnote(
    general = " ' *** ' : p < 0.001; ' ** ' : p < 0.01;
              ' * ' : p < 0.05; ' . ' : p < 0.1",
    general_title = "Significance Levels: ",
    footnote_as_chunk = TRUE
  ) %>%
  footnote(
    general = " '-' indicates not estimated",
    general_title = "Note: ",
    footnote_as_chunk = TRUE
  )
)

#Part 3
gof.choices<-control.gof.ergm(nsim=2000)
MOD3sim2000<-gof(MOD3, GOF=~model+idegree+odegree+distance+triadcensus,

```

```

        control=gof.choices)
MOD3sim2000$summary.model

par(mfrow = c(3, 3))
par(mar=c(3,3,3,3))
hist(MOD3sim2000$sim.model[,1]+.01,
     nclass=30,
     main = paste("Histogram of edges"),
     probability = T,
     xlab = NA)
abline(v = MOD3sim2000$summary.model[1,1], col = "red", lwd = 3)
abline(v = MOD3sim2000$summary.model[1,3], col = "blue", lwd = 3, lty=2)
hist(MOD3sim2000$sim.model[,2]+.01,
     nclass=30,
     main = paste("Histogram of mutual"),
     probability = T,
     xlab = NA)
abline(v = MOD3sim2000$summary.model[2,1], col = "red", lwd = 3)
abline(v = MOD3sim2000$summary.model[2,3], col = "blue", lwd = 3, lty=2)
hist(MOD3sim2000$sim.model[,3]+.01,
     nclass=30,
     main = paste("Histogram of gwideg"),
     probability = T,
     xlab = NA)
abline(v = MOD3sim2000$summary.model[3,1], col = "red", lwd = 3)
abline(v = MOD3sim2000$summary.model[3,3], col = "blue", lwd = 3, lty=2)
hist(MOD3sim2000$sim.model[,4]+.01,
     nclass=30,
     main = paste("Histogram of gwodeg"),
     probability = T,
     xlab = NA)
abline(v = MOD3sim2000$summary.model[4,1], col = "red", lwd = 3)
abline(v = MOD3sim2000$summary.model[4,3], col = "blue", lwd = 3, lty=2)
hist(MOD3sim2000$sim.model[,5]+.01,
     nclass=30,
     main = paste("Histogram of gwesp.OTP"),
     probability = T,
     xlab = NA)
abline(v = MOD3sim2000$summary.model[5,1], col = "red", lwd = 3)
abline(v = MOD3sim2000$summary.model[5,3], col = "blue", lwd = 3, lty=2)
hist(MOD3sim2000$sim.model[,6]+.01,
     nclass=30,
     main = paste("Histogram of gwesp.ITP"),

```

```

    probability = T,
    xlab = NA)
abline(v = MOD3sim2000$summary.model[6,1], col = "red", lwd = 3)
abline(v = MOD3sim2000$summary.model[6,3], col = "blue", lwd = 3, lty=2)
hist(MOD3sim2000$sim.model[,7]+.01,
     nclass=30,
     main = paste("Histogram of nodeicov.Gender"),
     probability = T,
     xlab = NA)
abline(v = MOD3sim2000$summary.model[7,1], col = "red", lwd = 3)
abline(v = MOD3sim2000$summary.model[7,3], col = "blue", lwd = 3, lty=2)
hist(MOD3sim2000$sim.model[,8]+.01,
     nclass=30,
     main = paste("Histogram of nodeocov.Gender"),
     probability = T,
     xlab = NA)
abline(v = MOD3sim2000$summary.model[8,1], col = "red", lwd = 3)
abline(v = MOD3sim2000$summary.model[8,3], col = "blue", lwd = 3, lty=2)
hist(MOD3sim2000$sim.model[,9]+.01,
     nclass=30,
     main = paste("Histogram of nodematch.Gender"),
     probability = T,
     xlab = NA)
abline(v = MOD3sim2000$summary.model[9,1], col = "red", lwd = 3)
abline(v = MOD3sim2000$summary.model[9,3], col = "blue", lwd = 3, lty=2)

par(mfrow=c(3,3))
plot(MOD3sim2000$sim.model[,1], type="l",
     main = paste("Trace plot for edges"), ylab="", xlab="")
plot(MOD3sim2000$sim.model[,2], type="l",
     main = paste("Trace plot for mutual"), ylab="", xlab="")
plot(MOD3sim2000$sim.model[,3], type="l",
     main = paste("Trace plot for gwideg"), ylab="", xlab="")
plot(MOD3sim2000$sim.model[,4], type="l",
     main = paste("Trace plot for gwodeg"), ylab="", xlab="")
plot(MOD3sim2000$sim.model[,5], type="l",
     main = paste("Trace plot for gwesp.0TP"), ylab="", xlab="")
plot(MOD3sim2000$sim.model[,6], type="l",
     main = paste("Trace plot for gwesp.ITP"), ylab="", xlab="")
plot(MOD3sim2000$sim.model[,7], type="l",
     main = paste("Trace plot for nodeicov"),
     ylab="", xlab="")
plot(MOD3sim2000$sim.model[,8], type="l",
     main = paste("Trace plot for nodeocov"),

```

```

      ylab="", xlab="")
plot(MOD3sim2000$sim.model[,9], type="l",
     main = paste("Trace plot for nodematch"),
     ylab="", xlab="")

plot(MOD3sim2000)

par(mfrow=c(4,2), mar=c(3,3,1,1))
for (k in 1:16)
{
  hist(MOD3sim2000$sim.triadecensus[,k],
       main=colnames(MOD3sim2000$sim.triadecensus)[k])

  abline(v = MOD3sim2000$obs.triadecensus[k],
         col = "blue", lwd = 3, lty=2)
}

```