**Assignment 14**

**Guan Yue Wang**

**Abstract**

The data is collected from subreddit ‘personalFinance’ to identify the network of the community

**Data Collection**

Reddit data is collected through Reddit API utilizing PRAW library in Python

**File Fomat**

Reddit comment is extracted from the web as author respondent pair and then converted to matrix

For example

Data is collected using reddit API command and saved as csv as below

Alice Tom

Alice Catherine

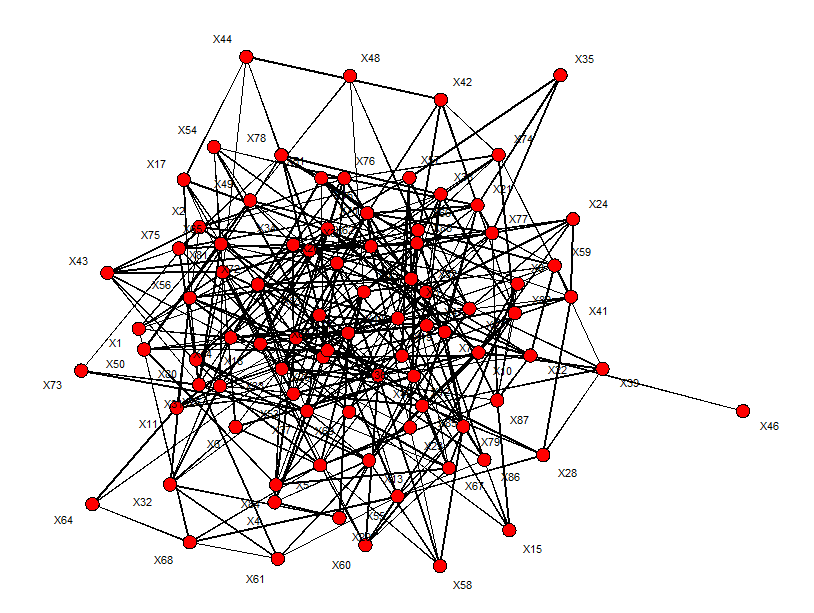
Catherine Alice

Then this is converted to Matrix as below

|  |  |  |  |
| --- | --- | --- | --- |
|  | Alice | Tom | Catherine |
| Alice | 0 | 1 | 1 |
| Tom | 1 | 0 | 0 |
| Catherine | 1 | 0 | 0 |

After that, this matrix would be loaded into R for further analysis

**ERGM**



==========================

Summary of model fit

==========================

Formula: ga.net ~ edges

Iterations: 6 out of 20

Monte Carlo MLE Results:

Estimate Std. Error MCMC % z value Pr(>|z|)

edges -2.33461 0.05697 0 -40.98 <1e-04 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Null Deviance: 5307 on 3828 degrees of freedom

Residual Deviance: 2286 on 3827 degrees of freedom

AIC: 2288 BIC: 2294 (Smaller is better.)

**RSiena on S50 Dataset Final Output**

Estimates, standard errors and convergence t-ratios

Estimate Standard Convergence

Error t-ratio

Rate parameters:

0.1 Rate parameter period 1 6.6421 ( 1.1556 )

0.2 Rate parameter period 2 5.1879 ( 0.9004 )

Other parameters:

1. eval outdegree (density) -2.7207 ( 0.1238 ) 0.0029

2. eval reciprocity 2.4384 ( 0.2104 ) 0.0116

3. eval transitive triplets 0.6299 ( 0.1392 ) -0.0158

4. eval 3-cycles -0.0561 ( 0.2665 ) 0.0105

5. eval smoke1 similarity 0.2615 ( 0.2042 ) 0.0011

6. eval alcohol alter -0.0206 ( 0.0742 ) 0.0240

7. eval alcohol ego 0.0369 ( 0.0783 ) 0.0201

8. eval alcohol ego x alcohol alter 0.1270 ( 0.0494 ) 0.0101

Overall maximum convergence ratio: 0.2285

Total of 2468 iteration steps.