Study of co-evolution of subreddit connections and anxiety

Priestley Fernandes

Abstract– Mental health issues play a significant role in deterioration of health among individuals. Reddit discussion is a popular source of data for studying online activities of users and identifying various trends based on behavior and information diffusion. This study is conducted to get insights about anxiety prevailing in an online network using the visualization tool of Gephi and statistical models like ERGM (Exponential Random Graph Model) and SAOM (Stochastic Actor Oriented Model) . The challenge is collapsing a network into a representative sample. The statistical models have certain limitations pertaining to the size of data. An interesting validation the study reveals is how anxiety-based discussions are leading to an increase in frequent connections among subreddits, which represents the users in it. Visualizations along with statistical modeling can decipher the intricacies of a real network and prove the extent of its utility as a powerful analytical and modeling tool.

I. Introduction

In the modern age of digitalization, the excessive dependency on online information and social networks has led to spread of misinformation. Consequently, free speech in the media has promoted sensitive discussions which eventually drives poor mental health. In *[1] Perera, G. et al. 2016,* a counterargument states that social networks help individuals communicate effectively and reduce anxiety. *[2] Gkotsis, G. et al 2017,* explains how Natural language processing of electronic health records is increasingly used to study mental health conditions and risk behaviors on a large scale. *[3] Bai W, Xi HT et al. 2023* present recent findings which suggest that online social connection buffered the negative effects of health anxiety under conditions of isolation during COVID-19.

Other studies have found a relationship between mental health disorders and increased usage of social networking platforms during COVID-19. As a whole, the SNS(Social Network Sites) environment may be just as complex as face-to-face interactions. As SNS membership continues to rise, it is becoming increasingly important to address the possible benefits and detriments the use of SNSs may have on mental health. Social network analysis is a popular tool gaining importance with time. This paper attempts to use this power to provide new ways of analyzing reddit connections through visualizations and modeling.

II. Motivation and research questions

[4] Zubair U et al. 2023 has a record which states that Mental health disorders have been a challenge since the invention of social media platforms. Social media is one of the major reasons for disability among psychiatric disorders. A great deal of literature has tried to establish links between social media exposure and mental health diseases. A literature presents critical perspectives on social media-induced psychiatric disorders to establish a holistic approach for their prevention and treatment. The use of social networks is strongly correlated with the development of anxiety and other psychological problems such as depression, insomnia, stress, decreased subjective happiness, and a sense of mental deprivation. A majority of research predicts that the likelihood of social media-induced mental health problems is directly proportional to the amount of time spent on these sites, the frequency of usage, and the number of platforms being used. These possible explanations include diffidence due to unhealthy comparisons, lack of emotional regulation due excessive exposure to social media and reduced real life social interactions. It is postulated that past anxiety leads to overuse of social media which indirectly serves as a coping strategy. Increased digitization led to a dynamic shift into a virtual world which has led to disastrous consequences on mental health of a significant population. This research is explored to answer a few questions using network structures existing in online communities of reddit.

Network theory has a unique feature of accounting for relationships between nodes. In this project, subreddits act as nodes and their interactions act as edges. In the absence of a network, modeling the relationship between connected nodes is not possible. A new way of exploring interactions between subreddits using exchange of anxious information leads a path to new inquiries and information. Using a combination of network visualization and statistical modeling, following questions are tackled:

1. How do anxiety related discussions propagate via the subreddit network?
2. Does the exchange of such posts have an impact on connections between subreddits?
3. How various network and dyadic attributes affect the co-evolution of interactions and exchange of anxiety related posts?

III. Background and related work

In *[5] Srijan Kumar et al. 2018, a* research group in Stanford developed a model which can predict whether a conflict is going to occur or not. A data driven view of community interactions and conflict is presented which makes use of deep learning and natural language processing to extract meaning from the data. . The preliminary data used in this project has been prepared as a part of this group. [6] Authors try to identify whether properties of the user- user network helps in predicting the virality of a post. The network structure allows capturing dependencies between users on the network.

In *[7] Kartik Sawhney et al.* , it is concluded that viral spreading from one local to another local community does not lead to a significant contribution to increasing the chances of diffusion events becoming more viral. The paper “N-GRAM Graphs for Topic extraction in educational forums", demonstrates that graph methods for computing centrality can be used to extract important topics and users along with deeper insights into the structure. Some papers explain a method to predict content characteristics based on network features. In paper, ‘Subrecommendit: Recommendation Systems on a Large-Scale Bipartite Graph’ explores community detection as a potential tool for recommendation.

In the paper, ‘It's complicated: A Visual Exploration of the Political Landscape of Reddit’, authors try to translate user behavior into relationships among subreddits, and also provide a troll detection algorithm to remove partially distracting contents. Social networks and the attributes of the actors in networks are not static. With time, they evolve interdependently. Stochastic Actor Oriented Model allows for statistical mechanisms of such a process. In *[8] Nynke et al. 2023*, authors study how stress changes in dynamic social conditions. Another group as per *[10] Joris J. Ebbers el at. 2019*, suggested a method for studying the co-evolution of social networks and individual actor behavior (SIENA). They studied how selection system orientations are core constituents of institutional logics affect, and are affected by, social networks.

IV. The Data

1. Data Description

Reddit is an American social news aggregation, content rating, and discussion website. Registered users submit content like links, text posts, images, and videos. These are then voted up and down by other members. Posts are organized by subject into user-created boards called “communities” or “subreddits”. Reddit administrators moderate communities, who are not reddit employees.

Data is taken from SNAP (Stanford Network Analysis Project). It is a huge dataset where records indicate hyperlinks between subreddits. It's a directed network, signed and temporal network containing almost 5 years data. Below is a tabular description of data.

Dataset Statistics:The data, as mentioned earlier, was created as described in [5] using natural language processing techniques.Some of the columns corresponding to the data are:

SOURCE_SUBREDDIT: the subreddit where the link originates

TARGET_SUBREDDIT: the subreddit where the link ends

POST_ID: the post in the source subreddit that starts the link

TIMESTAMP: time time of the post

POST_LABEL: label indicating if the source post is explicitly negative towards the target post. The value is -1 if the source is negative towards the target, and 1 if it is neutral or positive. The label is created using crowd-sourcing and training a text based classifier, and is better than simple sentiment analysis of the posts. Please see the reference paper for details.

POST_PROPERTIES: a vector representing the text properties of the source post, listed as a list of comma separated numbers. The vector elements include variables like

The post properties vector includes 87 attributes like: Number of characters , Number of characters without counting white space, Fraction of alphabetical characters, Fraction of digits, Fraction of uppercase characters, Fraction of white spaces, Fraction of special characters, such as comma, exclamation mark, etc., Number of words, Number of unique works,. Average number of words per sentence,Positive sentiment calculated by VADER, Negative sentiment calculated by VADER, Compound sentiment calculated by VADER, LIWC_Past, LIWC_Present, LIWC_Future, LIWC_Social, LIWC_Affect, LIWC_Anx, LIWC_Anger, LIWC_Sad, LIWC_Percept, LIWC_Sexual, LIWC_Ingest, etc. LIWC refers to Linguistic Inquiry and Word Count which gives a proportion of topic based words. For example, LIWC_Anx states the extent to which a particular post conveys anxiety. The same variable LIWC_Anx is the basis of analysis which will be considered as a network.

2. Data preparation and processing

2.1 Popular interactions

Dataset involved 5 time frames from 2014 to 2017 with final year having comparably lesser nodes. Since the dataset was huge, a filtering paradigm becomes a crucial consideration. Creation of samples which can be used to model data effectively becomes a priority. The entire dataset was loaded in google collab in a python enviroment. Most important features as per the purpose of study were LIWC_Anx and the frequency of interactions derived by grouping the edges. The following steps assured reliable data for drawing insights.

1. Finding missing values:

No missing values were encountered during analysis.

2. Removing duplicate records

There were no duplicates present in the dataset.

3. Finding popular interactions

A threshold of frequency was set as 25 based on a purpose of reducing overall number of nodes considering complex modeling involved in SAOM(Stochastic Actor Oriented Model) and ERGM (Exponential Random Graph Model). This reduced the overall size of data to 818 nodes. These nodes acted as important nodes for the selection process. This step later leads to a big hurdle in the model further which leads to a need for a better approach for modeling. This data is then converted into time frames of 2014, 2015, 2016 and 2017. The later years from 2015 to 2017 were selected to account for recent changes.

2.2 Connections dataset and Anxiety dataset

Interactions pertaining to 818 nodes were fetched from the main dataset by using a condition which pointed out the presence of one of the nodes in both source and target and number of interactions as more than 2 based on distribution plot. Many threshold values were experimented with, but a value of 2 led to a better insight. A threshold greater than 2, meant considerable interaction intensity(tie formed) while lesser meant insignificant interaction (No tie). At each step, the density was ensured below 0.2 to avoid complications in the model. This reduced the number of nodes to 658 nodes.

A similar process as described for connections dataset was used for creation of anxiety dataset. The only difference was the threshold of 0. A threshold greater than 0, meant a significant exchange of anxious posts while lesser meant less or insignificant interaction. Further, time frames played an important role to filter out the majority of relevant nodes.

2.3 Final sample

Final sample was created using all six subsets, each corresponding to a combination of connection/ anxiety with years 2015/2016/2017. Each subset was converted into a set of unique nodes using the set() method. Intersection of all these sets led to a list of 372 unique nodes. On applying a condition for checking whether every edge has a source and target in the unique node list, a final dataset for each time frame was created.

This approach led to an issue though. The unique nodes in each dataset were not 372. Few time frames like anxiety 2017 had two nodes missing. This posed a new alternative of either adding or eliminating nodes in order to maintain consistency across all time frames. Elimination would have led to previous inconsistency in the number of nodes across subsets since eliminating a node would have led to elimination of its interacting nodes completely. In order to address this challenge, a good alternative was to add nodes in subsets with missing nodes, such that the added node forms no ties with any of the nodes in the dataset. This manipulation was implemented keeping in mind that two nodes would not impact the results of modeling.

V. Methodology

Data preparation is an important aspect of this project which was covered in the previous section. The subsequent steps involved using the subsets of data to form insightful networks, analyzing the structure of networks, network visualization and modeling. Some of the assumptions made in this project are:

- Nodes remain the same across intervals
- Intensity of frequency and anxiety is same above threshold
- Time is continuous, although parameter estimation takes place at intervals
- Observed network is a consequence of previous network
- Impact of factors is neglected
- Actors have same propensity to change their ties

1. Network Analysis

R and Gephi are the most important tools used as a part of this project. For network analysis, the subsets were loaded and analyzed separately. The important features were source and target which were loaded using a csv file. Attributes like degree, modularity (Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, Fast unfolding of communities in large networks, in Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P1000), network diameter (Ulrik Brandes, A Faster Algorithm for Betweenness Centrality, in Journal of Mathematical Sociology 25(2):163-177, (2001)) and density were calculated. Results for each subset are presented in the results section.

2. Network Visualization

Gephi is the primary tool used for visualization. Gephi is an open source software which is used for manipulating, analyzing and visualizing large scale graph data. The following steps were used for visualizing the data for each network.
For each network, sizes of nodes were depicted using overall degree, in degree and outdegree. The color of nodes was representing the community associated with the node. Modularity was used as a partition filter in the colors section. Layouts like Fruchterman Reingold, ForceAtlas2 + Label Adjust were used for visualizations. Finally a moving timestamp video explaining the evolution of connections using an interval window of 1 sec.

3. Modeling

The aim was to study how an anxiety network was affecting connections among subreddits. To study various factors responsible for frequency of connections, the plan is to first study what factors affect it in a static environment using ERGM and then study the coevolution of the networks with time. This process of temporal analysis scans both the time variant and edge variant properties.

3.1 ERGM:

Exponential Random Graph Models (ERGMs) are a family of statistical models for analyzing data from social and other networks. Examples of networks examined using ERGM include knowledge networks, organizational networks, colleague networks, social media networks, networks of scientific development, and others.

The network data was converted into an adjacency matrix with indices corresponding to subreddit names. The rows were sorted using row names so that all matrices have the same node order across all time stamps. For the ERGM model, since temporal data cannot be used, the focus was on a single time frame. Year 2015 was studied using the model as below:

```
con_anx_2015_2 <- ergm(frequency_data_2015 ~ edges + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))
```

ergm() function is used for modeling the data.
Frequency_data_2015 - Dependent variable
Edges - Density of connection network edgecov(anxiety_data_2015_mat) - term used to capture the variation of anxiety ties (1 or 0) in 2015.
gwidegree(cutoff = 200) - In degree of subreddits

control.ergm() functions help in customizing the model to achieve better performance.
In order to improve convergence speed, parameters like parallel (helps in processing data across multiple CPU cores available) were used.
Reciprocity (mutual) , in-degree of nodes (gwidegree), influence parameters (gwesp), etc. were used to check for their significance.

Some cases involved checking the variation of connection ties with anxiety ties of previous year alone. On carrying out the same process for 2016, it was observed that past yearly ties of interaction network and anxiety ties do play a significant role in shaping the new ties. Thus, analysis proceeded with the year 2017 which allowed for checking all relevant past values. The below model provides better insights into the formation of ties in interaction networks.

```
ergm(formula = frequency_data_2017 ~ edges + gwidegree(cutoff = 200) +
            edgecov(anxiety_data_2015_mat) + edgecov(anxiety_data_2016_mat) +
            edgecov(anxiety_data_2017_mat) + edgecov(frequency_data_2015),
    control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))
```

3.2 SAOM (Stochastic Actor Oriented Model):

The stochastic actor oriented model (SAOM) is a method for modeling social interactions and social behavior over time. It can be used to model drivers of dynamic interactions using both exogenous covariates and endogenous network configurations, but also the co-evolution of behavior and social interactions.

The SAOM model was constructed using the Rsiena library in R. The following steps were used to create the model.

a) Formation of dependent network

Connections temporal network is an array of adjacency matrices of consecutive timestamps of 2015, 2016 and 2017. The entire array is used to create a siena dependent object.

```
connections <- array( c( frequency_data_2015, frequency_data_2016, frequency_data_2017),
              dim = c( 372, 372, 3 ) )


#Creation of Siena dependent variable
connections <- sienaDependent(connections)
```

b) Formation of varying dyad covariate

Similar to network connections, anxiety array consisting of time variant anxiety ties (1 or 0) is created. In this model, this array acts as an independent variable. The function varDyadCovar() is used to generate such an object.

```
anxiety <- array( c( anxiety_data_2015_mat, anxiety_data_2016_mat, anxiety_data_2017_mat),
            dim = c( 372, 372, 2 ))

#Creayion of varDyadCovar
anxiety_var <- varDyadCovar(anxiety)
```

c) Creation of Siena Data Object

The two objects are combined to create an object which can store both the network data and introduce plausible effects.

```
mydata <- sienaDataCreate(connections, anxiety_var)
```

d) Effects

Effects function is created to add or modify the different effects than can be introduced in the model. Density is the basic effect imposed.

```
myeff <- getEffects( mydata)
```

e) Siena model

Siena has three important inputs. The siena algorithm object, science data object and effects. The sienaAlgorithmCretae() function creates an algorithm of object while siena07() function is used to run the model.

```
myalgorithm <- sienaAlgorithmCreate( projname = 'connections_vs_anxiety',useStdInits = TRUE)

ans1 <- siena07( myalgorithm, data = mydata, effects = myeff)
```

VI.Results

The results section is divided as per methodologies or steps used for this project.

1. Network Analysis

1.1 Network 1: Connections/ Interactions

	2015	2016	2017
Density	0.009	0.01	0.06
Avg Degree	3.449	3.695	2.091
Modularity	0.503	0.553	0.751
Avg path length	4.668	5.962	1.876
Network diameter	15	17	6

From the above table it is clear that connection networks are sparse across time. The average degree in 2015 and 2016 is approximately 3 which is also a result of sampling methodology used. It indicates that overall the network is active. In 2017, a drop can be seen in connections. Subreddits like subredditdrama and askreddit are major nodes which can be clear from the high degree or interactions.

A higher modularity in the network across times leads to an inference that there is some implicit clustering which has led to formation of communities. The communities can be visible in the visualizations in the upcoming section. Average path length and network diameter increase from year 2015 to 2016 validates the expansion of a network which leads to higher complexity in navigation. Lower values for the same in 2017 are due to less data in 2017.

1.2 Network 2: Anxiety

	2015	2016	2017
Density	0.013	0.013	0.006
Avg Degree	4.847	4.97	2.22
Modularity	0.408	0.412	0.593
Avg path length	4.35	4.558	5.824
Network diameter	12	13	16

The anxiety network comparison with interaction networks using network statistics give some interesting insights. With an average degree of 5 in years 2015 and 2016, it can be seen that the interchange of anxious posts among subreddits is quite high and can be related to connections formed. Average path length is considerably higher too. Since it is comparable to the interaction degree, these networks seem to have some interdependency.

An interesting finding lies in the year 2017. In spite of the interactions being low due to less data variables, navigation is easier but anxiety propagation is complex. Somewhere the anxious posts in a smaller time frame take longer to propagate but which does not seem to impact the frequency of interactions. It can also be attributed to lower density of the anxiety network.

2. Network Visualization

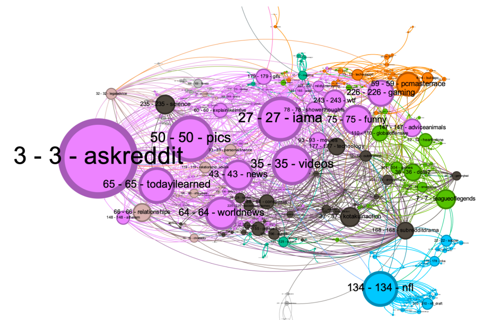
Across all time frames in interaction networks, it is quite evident that subreddit askreddit is the most popular while subredditdrama is the dominant influencer or active communicator. Major nodes formed in-degree distribution point towards a fact that popular nodes are information and entertainment sources on which the majority of users rely on for reliable information, news or fun. These include subreddits like askereddit, todayI learned, news, NFL (National Football League) , iama, etc. Segregation of communities can be seen wherein pink represents information and entertainment sources, blue represents sports or nfl related subredits, orange represents tech related support or tech help subrddits while green represents gaming communities. The above communities are as per 2015. SImilar trends can be observed in 2017 but color schemes differ.

In case of highly influential nodes, it can be observed that most of the subreddits are responsible for conveying opinions rather than facts. These sources spread unreliable information or are involved in morally depraved posts. One such subreddit is copypasta. This subreddit involves posts which are copy and pasted involving topics which are controversial or humorous. These are usually memes, but are highly influential since the majority of users follow them. The communities formed here are hatred spreading communities, conspiracy and drama, nfl and controversial topic talks.

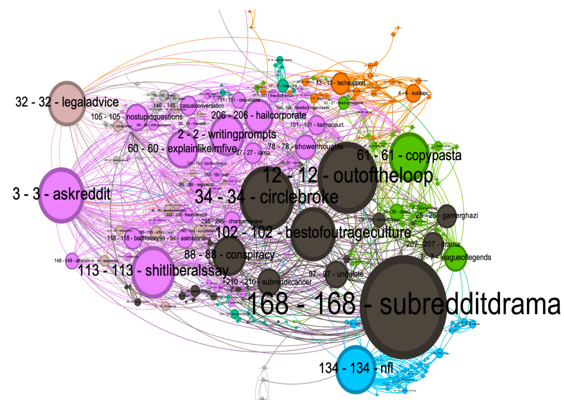
In 2017, smaller data leads to three clearly visible communities. One belongs to information sources, the other is gaming and the most widespread is NFL communities. For outdegree, majority of nodes belong to the drama community, but here too NFL has a good share with many subreddits within the community having higher outdegrees.

In case of an anxiety network, the communities will play an important role. Majority of users who are popular belong to gaming, information and entertainment sources, and advice related subreddits. Similar to interaction network, majority of influencer nodes belong to hatred spreading communities, conspiracy and drama, nfl and controversial topic talks. The new kind of subreddits which can be seen emerging in anxiety networks are advice related networks. All these belong to topics like relationships, law, property, depression, etc. These are all topics which contribute to anxiety issues.

2.1 Connection 2015



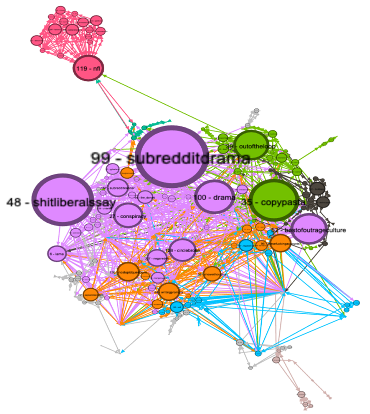
Indegree 2015



Outdegree 2015

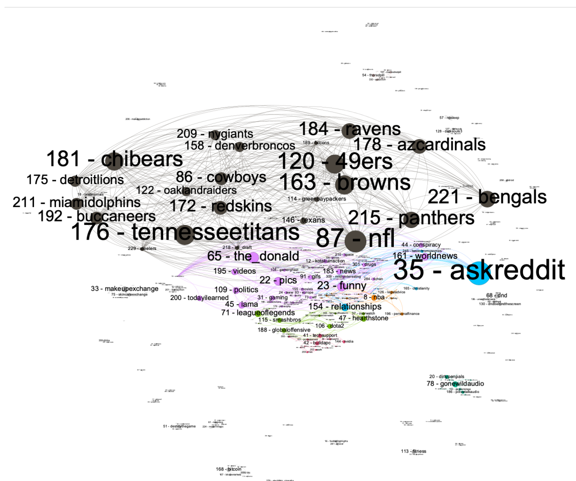
2.2 Connections 2016

Indegree 2016

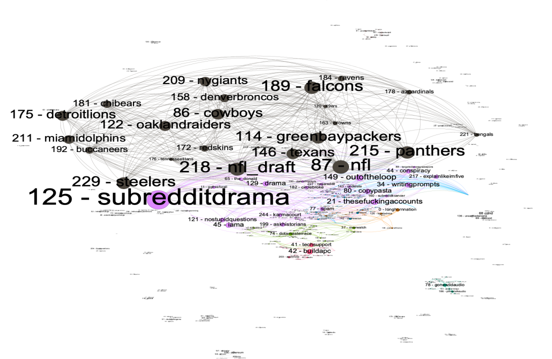


Outdegree 2016

2.3 Connections 2017

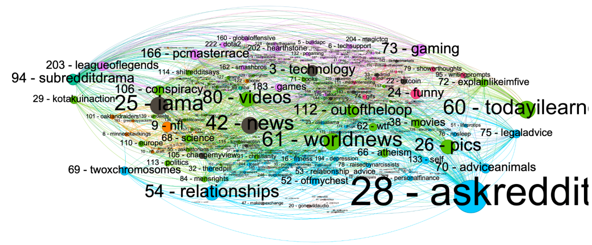


Indegree 2017

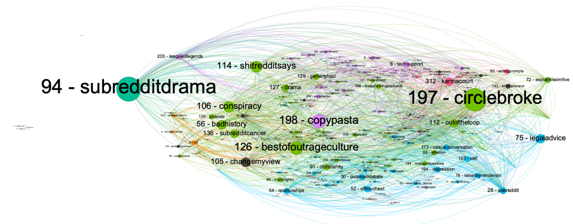


Outdegree 2017

2.4 Anxiety 2015

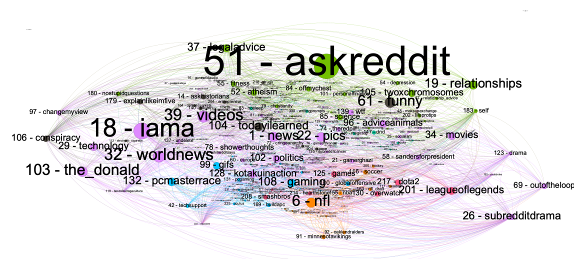


Indegree 2015

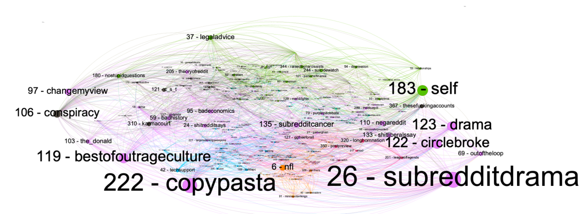


Outdegree 2015

2.4 Anxiety 2016

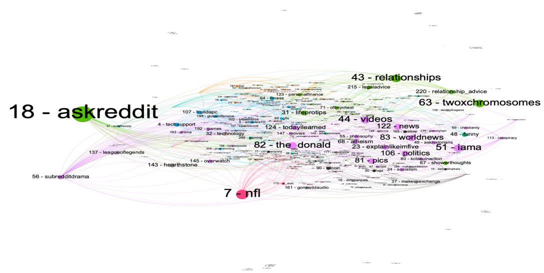


Indegree 2016

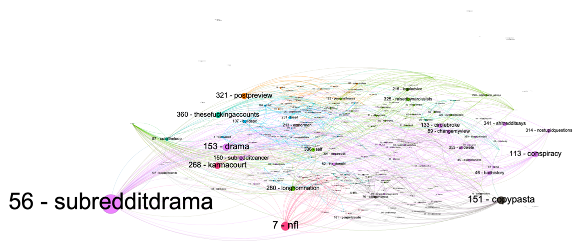


Outdegree 2016

2.6 Anxiety 2017



Indegree 2017



Outdegree 2017

3. Modeling:

3.1 ERGM

The results below are based on two models for the year 2015 alone. Terms for reciprocity (mutual), density(edges), indegree (gwidegree) , outdegree (gwodegree), variable for anxiety ties, etc. were tried one by one in the model. Terms related to reciprocity and outdegree proved to be insignificant with a p-value greater than 0.05. The indegree and transitivity (gwesp) parameters together led to insignificant contribution of gwesp.

The basic model is as follows:

Maximum Likelihood Results: (Year 2015)

	Estimate	Std. Error	MCMC %	z value	Pr(>|z|)
edges ***	-4.7163	0.0289	0	-163.20	<1e-04
Edgecov .anxiety_data_ 2015_mat	1.5383	0.1237	0	12.44	<1e-04

The following model was for built with transitive term:

Monte Carlo Maximum Likelihood Results:

	Estimate	Std. Error	MCMC %	z value	Pr(>|z|)
edges ***	-5.27722	0.05135	0	-102.776	<1e-04
gwesp.OTP .fixed.0.25	1.71021	0.21941	0	7.795	<1e-04
Edgecov. Anxiety_ data_2015_mat	0.94993	0.23633	0	4.020	<1e-04

The following model was for built with in-degree term:

Monte Carlo Maximum Likelihood Results:

	Estimate	Std. Error	MCMC %	z value	Pr(>|z|)
edges	-2.22639	0.07348	0	-30.30	<1e-04
gwidegree	-3.64433	0.06247	0	-58.34	<1e-04
Gwidegree. decay	2.48791	0.05487	0	45.34	<1e-04
Edgecov. Anxiety_ Data_2015_ mat	0.95609	0.09515	0	10.05	<1e-04

The final model results are as follows as follows:

Monte Carlo Maximum Likelihood Results: (Year 2017)

	Estimate	Std. Error	MCMC %	z value	Pr(>|z|)	
edges	-2.83005	0.09999	0	-28.303	< 1e-04	***
gwidegree	-3.55080	0.08727	0	-40.687	< 1e-04	***
gwidegree.decay	2.04258	0.07324	0	27.889	< 1e-04	***
Edgecov.anxiety _data_2015_mat	0.51540	0.17160	0	3.003	0.00267	**
Edgecov.anxiet y_data_2016_mat	0.81792	0.15330	0	5.335	< 1e-04	***
Edgecov.anxiety _data_2017_mat	0.59007	0.27220	0	2.168	0.03018	*
Edgecov.frequency _data_2015	1.19055	0.13604	0	8.752	< 1e-04	***

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

Null Deviance: 191325 on 138012 degrees of

freedom
Residual Deviance: 8630 on 138005 degrees of freedom

AIC: 8644 BIC: 8713 (Smaller is better. MC Std. Err. = 3.618)

The above model indicates that past time has an impact on the current tie formation.

3.2 SAOM

The basic configuration was run for SAOM, but it failed to converge. The maximum overall convergence ratio was 11.9042 which was way higher than the minimum required (0.25) for a model to be valid. Thus, temporal analysis using two networks did not give required insights. After thorough investigation, it was found that the change across connections network across two timestamps was too high due to which the model could not converge even based on the basic configuration.

Estimates, standard errors and convergence t-ratios

	Estimate	SCE	t-ratio
1. rate constant connections rate (period 1)	248.5962	(22.2247)	-8.1056
2. rate constant connections rate (period 2)	29.2439	(1.2785)	0.0219
3. eval outdegree (density)	-2.7661	(0.0244)	-0.0599
4. eval reciprocity	2.2907	(0.0454)	-0.2574

Overall maximum convergence ratio: 11.9042 x

Total of 2546 iteration steps.

Below is an investigation which validated the discrepancy and led to changes in the initial approach of preparing data compatible with the SAOM model. This approach will be handled in future with a robust criteria for sampling data.

Comparing two time frames of interactions.

```
Info.qap <-
qaptest(list(frequency_data_2015_net,frequency_data_2016_net), gcor, g1=1, g2=2, reps = 1000)
```

The results below show that there is a very small similarity between the two frames which have led to modeling failure.

QAP Test Results

Estimated p-values:
p(f(perm) >= f(d)): 0
p(f(perm) <= f(d)): 1

Test Diagnostics:
Test Value (f(d)): 0.02775863
Replications: 1000
Distribution Summary:

Min:	-0.007378654
1stQ:	-0.002031675
Med:	-0.000503967
Mean:	0.0001407258
3rdQ:	0.002551449
Max:	0.01553697

VII. Conclusion

The rigorous analysis led to an understanding of subreddit structure which disclose few intricate details. For the sample, the connection network evolves with time with expansion indicated by higher network diameters and number of edges while the first quarter of 2017 can see it shrinking since it was in its initial stages. Exchange of anxiety related information too evolves but lesser interactions do hinder its propagation which can be seen by the difference in their behaviors in the first quarter of 2017.

Visualizations convey that users of popular subreddits have a high exposure to depressing information since the popular subreddits who are interacting frequently are also bombarded with a lot of anxiety induced posts. In degree plays a vital role in connections along with exchange of anxiety. Users of subreddits which ask for advice on various issues fall prey to such negativity which is induced by the subreddits like subreddit drama and copypasta which are notorious for spreading unreliable information. These posts are majorly composed of opinions on serious controversial topics with minimal moderation of content. Trash talks, abusive language are very common in such posts. Misinformation is highly prevalent in these posts. NFL and gaming related communities too have a considerable share of anxiety related discussions which can impact the mental health of individuals. In short, popular and influencer nodes play a very significant role in spreading anxiety. If there is a considerable exchange of depression related posts between subreddits, they tend to form connections frequently. The quantification of this observation can be modeled using statistical models.

ERGM model infers that density of network, in degree of nodes, past connections along with current and past records of exchange of depressing posts define how ties are going to be formed in the network. Another way to put it is that , in the network all of the above parameters have significant presence. A negative estimate for density in a model indicates that a lower density favors such a network or this network has a lower density(sparse). More popular a particular subreddit, the higher are chances of it forming ties or having frequent interactions. Also, popularity plays a much prominent role compared to anxiety exchange. The higher the exchange of negativity (here anxiety), higher the chances of forming ties. It might not be that impactful but it does play a role. These connections depend on past records of anxiety related posts too, especially of the year before the current year. For example, 2017 was impacted more by 2016 than 2015 and 2017. Also, the current interactions depend on how frequent the connections were in the past.

Dominant sources of information are subreddits which promote hateful posts or uncensored information. From ERGM, it is evident that anxiety does play a significant role in tie formations or frequent interaction. It is a bad sign. A positive correlation can lead to a conclusion that sharing or receiving anxious posts can lead to frequent interactions. A low goodness of fit indicates lack of explainability.

VIII. Future Scope

Some accounts become inactive or are less frequent in particular years. The effect of node exit and entry can be studied and compared with our results. A sample of just 372 nodes of a bigger subset can lead to information loss. A better sampling technique can be created or algorithms can be optimized for large data handling. Subreddits behaviors can be incorporated based on the average of follower's attributes and later be used in the SAOM model to study the effect of behavior on the network and vice versa. Relationships between various behavioral, social and emotion specific attributes can be studied. The similarity of networks across time frames needs to be considered. A biggest flaw in this approach is data preparation to make data compatible with the SAOM model. A better filtering criteria can be implemented by considering smaller time stamps which can capture smaller changes.

IX.References

[1] Perera, G. et al. Cohort profile of the south london and maudsley nhs foundation trust biomedical research center (slam brc) case register: current status and recent enhancement of an electronic mental health record-derived data resource. BMJ open 6, e008721 (2016)

[2] Gkotsis, G., Oellrich, A., Velupillai, S. et al. Characterisation of mental health conditions in social media using Informed Deep Learning. Sci Rep 7, 45141 (2017). https://doi.org/10.1038/srep45141

[3] Bai W, Xi HT, Zhu Q, Ji M, Zhang H, Yang BX, Cai H, Liu R, Zhao YJ, Chen L, Ge ZM, Wang Z, Han L, Chen P, Liu S, Cheung T, Tang YL, Jackson T, An F, Xiang YT. Network analysis of anxiety and depressive symptoms among nursing students during the COVID-19 pandemic. J Affect Disord. 2021 Nov 1;294:753-760. doi: 10.1016/j.jad.2021.07.072. Epub 2021 Jul 22. Erratum in: J Affect Disord. 2023 May 1;328:355. PMID: 34375200; PMCID: PMC8433813.

[4] Zubair U, Khan MK, Albashari M. Link between excessive social media use and psychiatric disorders. Ann Med Surg (Lond). 2023 Mar 27;85(4):875-878. doi: 10.1097/MS9.0000000000000112. PMID: 37113864; PMCID: PMC10129173.

[5] Srijan Kumar, William L. Hamilton, Jure Leskovec, and Dan Jurafsky. Community Interaction and Conflict on the Web. The Web Conference (WWW). 2018.

[6] Noah Makow,Matt Millican, Chenyao Yu Stanford University, A Network Model for Reddit Post Virality Prediction

[7] Kartik Sawhney, Marcella Cindy Prasetio, Suvadi Paul, Community Detection Using Graph Structure and Semantic Understanding of Text

[8] NYNKE M. D. NIEZINK1 AND TOM A. B. SNIJDERS, CO-EVOLUTION OF SOCIAL NETWORKS AND CONTINUOUS ACTOR ATTRIBUTES

*[9] WenchaoYu† CharuC.Aggarwal‡ WeiWang† Modeling Coevolution Across Multiple Networks**

*[10] Joris J. Ebbers, Nachoem M. Wijnberg,
The co-evolution of social networks and selection system orientations as core constituents of institutional logics of future entrepreneurs at school,
Journal of Business Venturing,
Volume 34, Issue 3,
2019,
Pages 558-577,
ISSN 0883-9026,*

https://doi.org/10.1016/j.jbusvent.2018.12.005.

https://medium.datadriveninvestor.com/sentiment-analysis-of-stocks-from-financial-news-using-python-82ebdcefb638

https://sraf.nd.edu/loughranmcdonald-master-dictionary/

https://analyticsindiamag.com/convolutional-neural-network-image-classification-overview/

https://www.ibm.com/cloud/learn/recurrent-neural-networks

APPENDIX

Code: ERGM

```
#library(igraph) #you can use igraph if you prefer for data processing

setwd("/Users/priestleyfernandes/Downloads")

library(sna)

library(ergm)

library(tidyverse)


######################################################

##Read files


#Connections

g1 <- read.csv(file = "freq_com_new_2015_2.csv", sep = ",", header=TRUE) #graph modelling language

g2 <- read.csv(file = "freq_com_new_2016_2.csv", sep = ",", header=TRUE) #graph modelling language

g3 <- read.csv(file = "freq_com_new_2017_2.csv", sep = ",", header=TRUE) #graph modelling language


#Anxiety

a1 <- read.csv(file = "anxiety_com_new_2015_2.csv", sep = ",", header=TRUE) #graph modelling language

a2 <- read.csv(file = "anxiety_com_new_2016_2.csv", sep = ",", header=TRUE) #graph modelling language

a3 <- read.csv(file = "anxiety_com_new_2017_2.csv", sep = ",", header=TRUE) #graph modelling language


#Displaying the data

print(retrieval)

print(allocation)

######################################################

#View data

#Conversion into matrices

######################################################

#Connections

#2015

g1

g1_mat <- as.matrix(g1[,2:3])

g1_mat

mediag1<-graph_from_edgelist(g1_mat,directed=T)

g1_mat <- as.matrix(mediag1)

g1_mat

class(mediag1)

g1_mat_sorted <- g1_mat[order(rownames(g1_mat)),]

g1_mat_sorted


#2016

g2

nrow(g2)

g2_mat <- as.matrix(g2[,2:3])

g2_mat

mediag2<-graph_from_edgelist(g2_mat,directed=T)

mediag2 <- add_vertices(mediag2,1,name = 'badkarma')

V(mediag2)
```

```
g2_mat <- as.matrix(mediag2)
g2_mat

g2_mat_sorted <- g2_mat[order(rownames(g2_mat)),]
g2_mat_sorted

#2017
g3
g3_mat <- as.matrix(g3[,2:3])
g3_mat
mediag3<-graph_from_edgelist(g3_mat,directed=T)
mediag3 <- add_vertices(mediag3,1,name = 'badkarma')
write_graph(mediag3, 'connection_2017.gml', format = 'gml')
g3_mat <- as.matrix(mediag3)
g3_mat

g3_mat_sorted <- g3_mat[order(rownames(g3_mat)),]
g3_mat_sorted

#Anxiety
#2015

a1
a1_mat <- as.matrix(a1[,2:3])
a1_mat
mediaga1<-graph_from_edgelist(a1_mat,directed=T)
write_graph(mediaga1, 'anxiety_2015.gml', format = 'gml')
plot(mediaga1)

a1_mat <- as.matrix(mediaga1)
a1_mat

a1_mat_sorted <- a1_mat[order(rownames(a1_mat)),]
a1_mat_sorted

#2016
a2
a2_mat <- as.matrix(a2[,2:3])
a2_mat
mediaga2<-graph_from_edgelist(a2_mat,directed=T)
write_graph(mediaga2, 'anxiety_2016.gml', format = 'gml')
a2_mat <- as.matrix(mediaga2)
a2_mat

a2_mat_sorted <- a2_mat[order(rownames(a2_mat)),]
a2_mat_sorted

#2017
a3_mat <- as.matrix(a3[,2:3])
a3_mat
mediaga3<-graph_from_edgelist(a3_mat,directed=T)
mediaga3 <- add_vertices(mediaga3,2,name =
c('badkarma','sixers'))
write_graph(mediaga3, 'anxiety_2017.gml', format = 'gml')
a3_mat <- as.matrix(mediaga3)
a3_mat

a3_mat_sorted <- a3_mat[order(rownames(a3_mat)),]
a3_mat_sorted

#Creation on node sets for comparison
node_set_1 <- V(mediag1)
```

```
length(node_set_1)

node_set_2 <- V(mediag2)

length(node_set_2)

node_set_3 <- V(mediag3)

length(node_set_3)

node_set_4 <- V(mediaga1)

length(node_set_4)

node_set_5 <- V(mediaga2)

length(node_set_5)

node_set_6 <- V(mediaga3)

length(node_set_6)


#Noxde check

x <- Reduce(intersect, list(node_set_1, node_set_2, node_set_3, node_set_4, node_set_5, node_set_6))    # Identify common elements

# "A" "D"

length(x)


############################################################

#ERGM model


#Creation of vafriables

#Conection

frequency_data_2015 <- g1_mat_sorted

#frequency_data_2015_net <- as.network(frequency_data_2015)

frequency_data_2016 <- g2_mat_sorted

#frequency_data_2016_net <- as.network(frequency_data_2016)

frequency_data_2017 <- g3_mat_sorted

#frequency_data_2017_net <- as.network(frequency_data_2017)

frequency_data_2015

#Anxiety

anxiety_data_2015 <- a1_mat_sorted

anxiety_data_2016 <- a2_mat_sorted

anxiety_data_2017 <- a3_mat_sorted


#Converting into a compatible matrix object


frequency_data_2015 <- matrix(frequency_data_2015, nrow = 372, ncol = 372)

frequency_data_2016 <- matrix(frequency_data_2016, nrow = 372, ncol = 372)

frequency_data_2017 <- matrix(frequency_data_2017, nrow = 372, ncol = 372)


anxiety_data_2015_mat <- matrix(anxiety_data_2015, nrow = 372, ncol = 372)

anxiety_data_2016_mat <- matrix(anxiety_data_2016, nrow = 372, ncol = 372)

anxiety_data_2017_mat <- matrix(anxiety_data_2017, nrow = 372, ncol = 372)


dim

####################################################


##ERGM Analysis


###2015


library(parallel)

detectCores()

#ERGM ANALYSIS


############
```

```
#2015

con_anx_2015 <- ergm(frequency_data_2015 ~ edges + gwidegree(cutoff = 200) + gwesp(0.25, fixed = T) + edgecov(anxiety_data_2015_mat), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2015

coef(con_anx_2015)

summary(con_anx_2015)



con_anx_2015_1 <- ergm(frequency_data_2015 ~ edges + edgecov(anxiety_data_2015_mat), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2015_1

coef(con_anx_2015_1)

summary(con_anx_2015_1)



gof_con_anx_2015_1 <- gof(con_anx_2015_1)

gof_con_anx_2015_1



con_anx_2015_2 <- ergm(frequency_data_2015 ~ edges + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2015_2

summary(con_anx_2015_2)



gof_con_anx_2015_2 <- gof(con_anx_2015_2)

gof_con_anx_2015_2



con_anx_2015_3 <- ergm(frequency_data_2015 ~ edges + gwesp(0.25, fixed = T) + edgecov(anxiety_data_2015_mat), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2015_3

summary(con_anx_2015_3)



gof_con_anx_2015_3 <- gof(con_anx_2015_3)

gof_con_anx_2015_3



##################

#2016

con_anx_2016_2_1 <- ergm(frequency_data_2016 ~ edges + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2015_2

summary(con_anx_2015_2)



#########################

##Analysis of 2017 data



con_anx_2017_2_1 <- ergm(frequency_data_2017 ~ edges + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat) + edgecov(anxiety_data_2016_mat), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2017_2_1

summary(con_anx_2017_2_1)



con_anx_2017_2_2 <- ergm(frequency_data_2017 ~ edges + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat) + edgecov(anxiety_data_2016_mat) + edgecov(anxiety_data_2017_mat), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2017_2_2

summary(con_anx_2017_2_2)
```

```
con_anx_2017_2_3 <- ergm(frequency_data_2017 ~ edges + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat) + edgecov(anxiety_data_2016_mat) + edgecov(anxiety_data_2017_mat) + edgecov(frequency_data_2016) , control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2017_2_3

summary(con_anx_2017_2_3)

con_anx_2017_2_4 <- ergm(frequency_data_2017 ~ edges + + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat) + edgecov(anxiety_data_2016_mat) + edgecov(anxiety_data_2017_mat) + edgecov(frequency_data_2016) + edgecov(frequency_data_2015), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2017_2_4

summary(con_anx_2017_2_4)

gofcon_anx_2017_2_4 <- gof(con_anx_2017_2_4)

gofcon_anx_2017_2_4

con_anx_2017_2_5 <- ergm(frequency_data_2017 ~ edges + + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat) + edgecov(anxiety_data_2016_mat) + edgecov(anxiety_data_2017_mat) + edgecov(frequency_data_2016) + edgecov(frequency_data_2015), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2017_2_5

summary(con_anx_2017_2_5)

gof_con_anx_2017_2_5 <- gof(con_anx_2017_2_5)

gof_con_anx_2017_2_5

con_anx_2017_2_6 <- ergm(frequency_data_2017 ~ edges + gwesp(0.25, fixed = T) + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat) + edgecov(anxiety_data_2016_mat) + edgecov(anxiety_data_2017_mat) + edgecov(frequency_data_2015) + edgecov(frequency_data_2016), control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2017_2_6

summary(con_anx_2017_2_6)

con_anx_2017_2_7 <- ergm(frequency_data_2017 ~ edges + gwidegree(cutoff = 200) + edgecov(anxiety_data_2015_mat) + edgecov(anxiety_data_2016_mat) + edgecov(anxiety_data_2017_mat) + edgecov(frequency_data_2015) , control = control.ergm(seed = 1, parallel = 6, parallel.type = "PSOCK"))

con_anx_2017_2_7

summary(con_anx_2017_2_7)

gof_con_anx_2017_2_7 <- gof(con_anx_2017_2_7)

summary(gof_con_anx_2017_2_7)

gof_con_anx_2017_2_7

par(mfrow=c(3,2))

par(mar = c(1, 1, 1, 1))

plot(gof_con_anx_2017_2_7)
```

Code: SAOM

```
install.packages("RSiena")

library("RSiena")

library("network")

library( sna )

install.packages("igraph")

library(igraph)

install.packages("tidyverse")
```

```
library(tidyverse)

install.packages("spam")

library("spam")


####################################

getwd()

setwd('/Users/priestleyfernandes/Downloads')


######################################



##Read files


#Connections

g1 <- read.csv(file = "freq_com_new_2015_2.csv", sep = ",", header=TRUE) #graph modelling language

g2 <- read.csv(file = "freq_com_new_2016_2.csv", sep = ",", header=TRUE) #graph modelling language

g3 <- read.csv(file = "freq_com_new_2017_2.csv", sep = ",", header=TRUE) #graph modelling language


#Anxiety

a1 <- read.csv(file = "anxiety_com_new_2015_2.csv", sep = ",", header=TRUE) #graph modelling language

a2 <- read.csv(file = "anxiety_com_new_2016_2.csv", sep = ",", header=TRUE) #graph modelling language

a3 <- read.csv(file = "anxiety_com_new_2017_2.csv", sep = ",", header=TRUE) #graph modelling language



#View data

#Conversion into matrices


#Connections

#2015

g1

g1_mat <- as.matrix(g1[,2:3])

g1_mat

mediag1<-graph_from_edgelist(g1_mat,directed=T)

g1_mat <- as.matrix(mediag1)

g1_mat

class(mediag1)

g1_mat_sorted <- g1_mat[order(rownames(g1_mat)),]

g1_mat_sorted



#2016

g2

nrow(g2)

g2_mat <- as.matrix(g2[,2:3])

g2_mat

mediag2<-graph_from_edgelist(g2_mat,directed=T)

mediag2 <- add_vertices(mediag2,1,name = 'badkarma')

V(mediag2)

g2_mat <- as.matrix(mediag2)

g2_mat


g2_mat_sorted <- g2_mat[order(rownames(g2_mat)),]

g2_mat_sorted


#2017

g3

g3_mat <- as.matrix(g3[,2:3])

g3_mat

mediag3<-graph_from_edgelist(g3_mat,directed=T)

mediag3 <- add_vertices(mediag3,1,name = 'badkarma')
```

```r
write_graph(mediag3, 'connection_2017.gml', format = 'gml')

g3_mat <- as.matrix(mediag3)

g3_mat


g3_mat_sorted <- g3_mat[order(rownames(g3_mat)),]

g3_mat_sorted


#Anxiety

#2015


a1

a1_mat <- as.matrix(a1[,2:3])

a1_mat

mediaga1<-graph_from_edgelist(a1_mat,directed=T)

write_graph(mediaga1, 'anxiety_2015.gml', format = 'gml')

plot(mediaga1)


a1_mat <- as.matrix(mediaga1)

a1_mat


a1_mat_sorted <- a1_mat[order(rownames(a1_mat)),]

a1_mat_sorted


#2016

a2

a2_mat <- as.matrix(a2[,2:3])

a2_mat

mediaga2<-graph_from_edgelist(a2_mat,directed=T)

write_graph(mediaga2, 'anxiety_2016.gml', format = 'gml')

a2_mat <- as.matrix(mediaga2)

a2_mat

a2_mat_sorted <- a2_mat[order(rownames(a2_mat)),]

a2_mat_sorted


#2017

a3_mat <- as.matrix(a3[,2:3])

a3_mat

mediaga3<-graph_from_edgelist(a3_mat,directed=T)

mediaga3 <- add_vertices(mediaga3,2,name =
c('badkarma','sixers'))

write_graph(mediaga3, 'anxiety_2017.gml', format = 'gml')

a3_mat <- as.matrix(mediaga3)

a3_mat


a3_mat_sorted <- a3_mat[order(rownames(a3_mat)),]

a3_mat_sorted


frequency_data_2015 <- g1_mat_sorted

frequency_data_2015_net <- as.network(frequency_data_2015)

frequency_data_2016 <- g2_mat_sorted

frequency_data_2016_net <- as.network(frequency_data_2016)

frequency_data_2017 <- g3_mat_sorted

frequency_data_2017_net <- as.network(frequency_data_2017)

frequency_data_2015

#Anxiety

anxiety_data_2015 <- a1_mat_sorted

anxiety_data_2016 <- a2_mat_sorted

anxiety_data_2017 <- a3_mat_sorted

#Creation on node sets for comparison

node_set_1 <- V(mediag1)

length(node_set_1)
```

```
node_set_2 <- V(mediag2)
length(node_set_2)
node_set_3 <- V(mediag3)
length(node_set_3)
node_set_4 <- V(mediaga1)
length(node_set_4)
node_set_5 <- V(mediaga2)
length(node_set_5)
node_set_6 <- V(mediaga3)
length(node_set_6)

info.qap <-
qaptest(list(frequency_data_2015_net,frequency_data_2016_net),
gcor, g1=1, g2=2, reps = 1000)

summary(info.qap)
info.qap1 <-
qaptest(list(frequency_data_2016_net,frequency_data_2017_net),
gcor, g1=1, g2=2, reps = 1000)
info.qap
info.qap1
summary(info.qap1)
summary(info.qap)
#Noxde check
x <- Reduce(intersect, list(node_set_1, node_set_2, node_set_3,
node_set_4, node_set_5, node_set_6))     # Identify common
elements
# "A" "D"
length(x)

#Siena Model

#Creation of vafriables
#Conection
frequency_data_2015 <- g1_mat_sorted
#frequency_data_2015_net <- as.network(frequency_data_2015)
frequency_data_2016 <- g2_mat_sorted
#frequency_data_2016_net <- as.network(frequency_data_2016)
frequency_data_2017 <- g3_mat_sorted
#frequency_data_2017_net <- as.network(frequency_data_2017)
frequency_data_2015
#Anxiety
anxiety_data_2015 <- a1_mat_sorted
anxiety_data_2016 <- a2_mat_sorted
anxiety_data_2017 <- a3_mat_sorted

class(frequency_data_2015)
class(frequency_data_2016)
class(frequency_data_2017)
dim(frequency_data_2015)
dim(frequency_data_2016)
dim(frequency_data_2017)

#Converting into a compatible matrix object

frequency_data_2015 <- matrix(frequency_data_2015, nrow =
372, ncol = 372)
frequency_data_2016 <- matrix(frequency_data_2016, nrow =
372, ncol = 372)
frequency_data_2017 <- matrix(frequency_data_2017, nrow =
372, ncol = 372)

anxiety_data_2015_mat <- matrix(anxiety_data_2015, nrow =
372, ncol = 372)
```

```
anxiety_data_2016_mat <- matrix(anxiety_data_2016, nrow = 372, ncol = 372)

anxiety_data_2017_mat <- matrix(anxiety_data_2017, nrow = 372, ncol = 372)

frequency_data_2015

frequency_data_2016

frequency_data_2017

anxiety_data_2015_mat

class(anxiety_data_2015)

class(frequency_data_2015)

# Dependent variable

#####################################################
#

#Connection Network: The dependent network

#Dimension 372 nodes

connections <- array( c( frequency_data_2015, frequency_data_2016, frequency_data_2017),

            dim = c( 372, 372, 3 ) )

connections

#Creation of Siena dependent variable

connections <- sienaDependent(connections)

dim(connections)

str(connections)

#Network 2: The Anxiety network as Varying Dyad Covariate

anxiety <- array( c( anxiety_data_2015_mat, anxiety_data_2016_mat, anxiety_data_2017_mat),

            dim = c( 372, 372, 2 ))

#Creayion of varDyadCovar

anxiety_var <- varDyadCovar(anxiety)

#Cross Check

anxiety_var

connections

#####################################################

#Creation of Siena Data by combing the two

mydata <- sienaDataCreate(connections, anxiety_var)

mydata

print01Report( mydata, modelname="connectionVsanxiety")

#####################################################

#Effects

myeff <- getEffects( mydata)

effectsDocumentation(myeff)

#myeff <- setEffect(myeff, recip , include = FALSE)

#myeff

myeff <- setEffect(myeff, density , include = TRUE)

myeff

#Inclusion of effects
```

?includeEffects

myeff <- includeEffects(myeff,outPop)

myeff

#Crreate Algorithm

?sienaAlgorithmCreate

myalgorithm <- sienaAlgorithmCreate(projname = 'connections_vs_anxiety',useStdInits = TRUE)

ans1 <- siena07(myalgorithm, data = mydata, effects = myeff)

ans1

summary(ans1)

ans1$tconv.max

ans <- siena07(myalgorithm, data = mydata, effects = myeff, prevAns=ans1)

ans

summary(ans)