Spring 6-9-2023

Information diffusion in online social networks: a simulation experiment

Maxwell Jacob Blum
maxwell.j.blum.23@dartmouth.edu

Information diffusion in online social networks: a simulation experiment

Maxwell Jacob Blum

Honors Thesis

Peter DeWan, Advisor

Program in Quantitative Social Science

Dartmouth College

June 4, 2023

Contents

List of Figures

List of Tables

1 Acknowledgements

When I first enrolled at Dartmouth College in September of 2019, I never thought I would have been able to write the pages that follow. In addition to partially fulfilling the Quantitative Social Science major, my hope for this honors thesis is that it can serve as a testament to the person I have become through my time at Dartmouth.

The people that helped make this project possible did not just help me produce what follows – they provided me with a set of tools that I believe will prove invaluable as I continue making decisions inspired by my love of learning. That said, I feel as though this project would be incomplete without extending my sincere thanks to a few individuals.

I would like to thank Professors Tracy Sweet and Hirokazu Shirado from the University of Maryland and Carnegie Mellon University, respectively, for taking the time to meet with a complete stranger to discuss his project.

In addition, I would like to thank my advisor, Professor Peter DeWan, for his encouragement and expertise.

Lastly, thank you to my family and friends for their endless support throughout this yearlong academic adventure.

2 Abstract

The advent of online social networks has completely transformed the way we communicate, with news, opinions, and ideas now spreading faster than ever before (Guille et al., 2013; Lee et al., 2022). That online social networks have a profound impact on the spread of information suggests further investigation of the relationship between network structure and information diffusion (Light & Moody, 2020).

This honors thesis investigates degree assortativity – a measure of large-scale network structure that has often only been a footnote in relevant literature on information diffusion in online social networks – and its effect on the speed of information diffusion in online social networks. Two rewiring algorithms (Xulvi-Brunet & Sokolov, 2005) were applied to rewire a Facebook friend circle ($n = 44$) with varying degree assortativity, ranging from approximately -0.7 to 0.4. For each of the 160 rewired graphs, a random node was selected to infect (i.e., spread information to) its neighbors with probabilities ranging from 10 to 50 percent, and the number of infected nodes after each round of diffusion was recorded.

Results suggest that degree assortativity and the speed of information diffusion have a strong inverse relationship – disassortative networks spread the same information faster. Moreover, degree assortativity appears to drive the speed of information diffusion more than its correlates, clustering coefficient and average path length (Xulvi-Brunet & Sokolov, 2005).

3 Introduction

Human social networks affect our lives, health, desires, feelings, thoughts, and actions (Yale Institute for Network Science, 2022). The study of the structure and function of social networks – relationships between people – can help us develop healthier, more cooperative, and smarter communities. Though thinking in terms of social networks dates to as early as 1908 (Furht, 2010; Simmel, 1908), social network analysis remains a largely underdeveloped arena compared to other sociological subfields.

One subtopic within this burgeoning field concerning social relationships is diffusion – the process through which people spread behaviors, infections, objects, and information through both simple and complex processes of contagion. While what exactly a person passes through their social network helps determine how it diffuses, so does the structure of their social network writ large. The interaction between structure and diffusion within social network analysis is not well-documented and warrants further investigation – exploring "linkages between structure and diffusion is a necessary next step in the growing body of research [of] experimental diffusion manipulations" (Light & Moody, 2020).

A noteworthy inhibitor to the growth of the field of social network analysis is the (in)feasibility of experimentally manipulating relationships between people (Light & Moody, 2020) – for instance, it is impossible to make two people become friends as an experimental treatment. Due to the innately low feasibility of social network experiments, simulation experiments are quite common, and their methods and results are often extrapolated to predict what one might expect if the simulated phenomenon were to occur in the real world (Mutlu & Garibay, 2020). Here, the power of a simulation is harnessed to hopefully illuminate the poorly understood relationship between structure and diffusion. Moreover, the promising results described in subsequent sections could potentially indicate to more sophisticated and endowed research groups that this topic is worthy of a more elaborate experiment.

This thesis aims to make a significant and specific stride in the field of social network analysis by investigating the relationship between network structure and diffusion. I experimentally manipulate degree assortativity – an important characteristic of a network's shape that, due to social network analysis being on the frontier of knowledge, is not well documented. I define degree assortativity, contextualize what is known about it, and test its effect on the speed of information diffusion through online social networks, a relationship that is yet to be tested. Namely, this thesis asks, does degree assortativity affect the rate at which information diffuses throughout simulated online social networks?

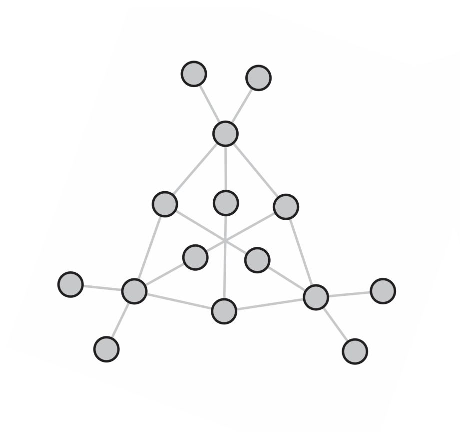
4 Literature Review

4.1 Degree assortativity

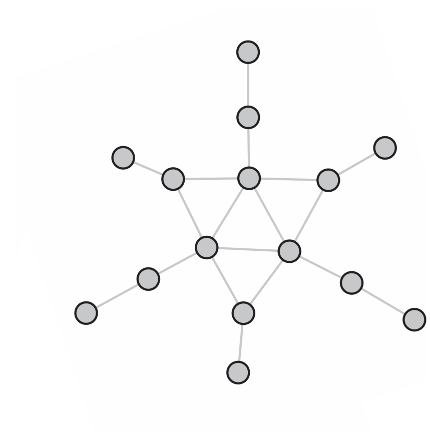
Broadly speaking, assortativity is a measure of a network's structure. Within the field of social network analysis, much attention is often disproportionately brought to properties of individual vertices – for example, "how far apart they are, what their degrees are, and so forth" (i.e., path length, number of connections, centrality) (Newman & Girvan, 2002). Though understanding the properties of individual vertices can often help identify who the most influential players are in a network, measures of large-scale network structure – such as assortativity – can help detect communities (Newman & Girvan, 2002), understand processes of threshold models of contagion (Mutlu & Garibay, 2020), and investigate how things and ideas flow through networks writ large (Vega-Oliveros et al., 2020).

Assortativity, or assortative mixing, is a measure of large-scale network structure that describes the extent to which 'birds of a feather flock together.' Newman and Girvan pioneered the concept of assortativity in 2003, describing assortativity as the "preferential association of network vertices with others that are like them in some way" (Newman & Girvan, 2002). Identifying in what way network vertices are similar to each other is crucial to any experimental design of assortativity. Although in theory the term assortativity can refer to any quality which nodes share, since at least 2005, social network analysis practitioners have tended to interchangeably address assortativity and degree assortativity (Xulvi-Brunet & Sokolov, 2005) – that is, the propensity for people to connect with other people with a similar degree (i.e., number of connections) (G. Wang et al., 2014). Broadly, degree assortatively has rarely been experimentally manipulated as an independent variable. (Mutlu & Garibay, 2020).

Figure 1 illustrates two networks designed by Shirado et al., 2019. While both networks contain fifteen nodes and eighteen edges, they vary in assortativity. In addition, both networks feature the same degree distribution – "6 nodes with 1 edge, 3 nodes with 2 edges, 3 nodes with 3 edges, and 3 nodes with 5 edges" (Shirado et al., 2019). Because both networks feature the same degree distribution, the point of this figure is to demonstrate how degree assortativity is a property of the connections between people, not the people themselves. Two distinct circles of friends can feature the same group of people with the same personal attributes, but a few rewired connections can impact how popular nodes and less popular nodes choose to connect with one another.



(a) Degree-disassortative network



(b) Degree-assortative network

Figure 1: Illustration of two networks with varying assortativity (Shirado et al., 2019, p. 2)

Assortativity, first quantified by Newman and Girvan, 2002, is measured as a Pearson correlation coefficient (Bliss et al., 2012; Hu & Wang, 2009).

> "As with assortative mixing on discrete characteristics, one can define an assortativity coefficient to quantify the extent to which mixing is biased according to scalar vertex properties. To do this, we define e_{xy} to be the fraction of edges in our network that connect a vertex of property x to another of property y. The matrix e_{xy} must satisfy sum rules as before, of the form

$$\sum_{xy} e_{xy} = 1, \ \sum_{y} e_{xy} = a_x, \sum_{x} e_{xy} = b_y, \tag{1}$$

where a_x and b_y are, respectively, the fraction of edges that start and end at vertices with ages x and y. Then the appropriate definition for the assortativity coefficient is

$$r = \frac{\sum_{xy} xy(e_{xy} - a_x b_y)}{\sigma_a \sigma_b}, \tag{2}$$

where σ_a and σ_b are the standard deviations of the distributions a_x and b_y." (Newman & Girvan, 2002)

The coefficient of degree assortativity, r, always ranges between -1 (disassortative) and 1 (assortative). A coefficient of degree assortativity of 0 is considered neutral or uncorrelated. For illustrative purposes, using the Pearson correlation coefficient to calculate the coefficient of degree assortativity for the degree-disassortative network in Figure 1, we find $r = -0.744$, and for the degree-assortative network, the coefficient of degree assortativity is $r = 0.233$ (Shirado et al., 2019). As evidenced in the methods section, the probability of obtaining strongly degree-assortative and degree-disassortative networks through random rewiring decreases as the number of nodes increases, so degree assortativity of magnitude $r > 0.3$ is rarely observed in large human social networks (Newman, 2001).

4.2 Different networks have different assortativities

Simply put, different social networks have different degree assortativities. "Many human networks are exceptionally degree assortative" (Newman & Girvan, 2002; Shirado et al., 2019). Numerous observational studies have measured degree assortativity as a dependent variable, and their findings suggest that social networks range in assortativity, both among and between online and offline social networks (Bliss et al., 2012; Cero & Witte, 2020; Hu & Wang, 2009; Mutlu & Garibay, 2020; Newman & Girvan, 2002; G. Wang et al., 2014). Broadly, there appears to be a disparity in assortativity between online versus offline social networks. While it is a popular hypothesis that social networks tend to be assortative (Mutlu & Garibay, 2020), the study of online social networks has challenged this hypothesis, as "many [online social networks] show [a] disassortative or neutral mixing feature" (Hu & Wang, 2009).

Table 1: Degree assortativity coefficients, r, across online and real-life social networks

Type	Network	r	Reference
Online social network	Facebook	0.116	G. Wang et al., 2014
	Twitter	-0.025	G. Wang et al., 2014
	Internet	-0.189	Qian Chen et al., 2002
	YouTube	-0.033	Mislove et al., 2007
	Reddit	-0.111	Mutlu and Garibay, 2020
	Instagram	-0.097	Ferrara et al., 2014
	LinkedIn	-0.199	Rajendran, 2021
	Whisper	-0.011	G. Wang et al., 2014
Real social network	Email address books	0.092	Newman et al., 2002
	Film actor collaborations	0.208	Watts and Strogatz, 1998
	Physics coauthorship	0.363	Newman, 2001

As shown in Table 1, degree assortativity can widely vary across online and offline human social networks. The focus of this paper is on assortativity in online social networks, and among them, Facebook stands alone in being degree-assortative.

4.2.1 The emergence of degree assortativity

That most online social networks are degree-disassortative calls into question the influence that high-degree individuals have on the assortativity of an online social network (Barabási & Bonabeau, 2003). Indeed, for most online social networks, high-degree individuals have a defining impact on the degree assortativity of the overall online social network, but as the degree assortativity of Facebook stands alone, so might the impact of high-degree individuals on the assortativity of Facebook. For instance, the impact of high-degree users on degree assortativity is similar for Twitter and Instagram but different for Facebook, and this is likely because the social role of high-degree individuals differs between Twitter and Instagram versus Facebook.

Twitter is characterized by its public and open nature in which high-degree individuals act as broadcasters, thus reaching a high number of users with their Tweets (Kwak et al., 2010). That the users with the highest degrees on Twitter tend to use the platform more as a site of broadcasting news media than as a social network likely drives, at least in part, the deviation of Twitter from the well-known trend of offline human social networks to be degree-assortative (Kwak et al., 2010; G. Wang et al., 2014). Instagram is a similar site of mass communication from the high-degree users to the low-degree users, but with its emphasis on visual content, the high-degree individuals of Instagram are particularly well-positioned to be successful social media marketers through posting simple, popular, visually-striking advertisements (Jaakonmäki et al., 2017). As such, Instagram has become an increasingly popular destination for marketers looking to grow their social media exposure (Jaakonmäki et al., 2017).

On another hand, because Facebook is more oriented toward personal connections and friendships that other online social networks, perhaps Facebook's high-degree individuals have a distinct impact on degree assortativity compared to high-degree individuals on other online social networks. Compared to Twitter and Instagram, friendships on Facebook tend to be longer-lasting, and Ellison et al., 2007 suggest that the greater similarities between friendships on Facebook and friendships in real life (compared to less similarities between friendships on other online social networks and real-life friendships) bridge online popularity with offline social capital. Therefore, high-degree individuals on Facebook might leverage their connections for offline social capital gain, whereas high-degree individuals on Twitter and Instagram might leverage their connections for online influencing, marketing, or broadcasting (Ellison et al., 2007; Jaakonmäki et al., 2017; Kwak et al., 2010). Thusly, because the self-presentation of a Facebook user is often more reflective of their offline self-presentation (Marwick, 2015), perhaps high-degree Facebook users have a lesser impact on the assortativity of the overall Facebook network than marketers on Instagram or broadcasters on Twitter.

In sum, Facebook is unique among other online social networks not only for its positive degree assortativity, but for its distinct connection with offline social capital. Simply put, high-degree individuals on Facebook might have high degrees due to the well-known tendency for humans to form offline connections assortatively (Kwak et al., 2010; Newman et al., 2002), whereas the same is not necessarily true for other online social networks. Thus, Facebook is certainly a special network for its positive degree assortativity and the online-offline similarities for high-degree individuals on the platform (Kwak et al., 2010; Newman et al., 2002; Ugander et al., 2011). Testing the effect of degree assortativity on the speed of information diffusion on Facebook could have profound implications on the digitization of human connection because the degree assortativity of Facebook might be more indicative of the degree assortativity of offline connections. Therefore, the effect of Facebook's degree assortativity on Facebook's speed of information diffusion could substantiate claims about the effect of offline degree assortativity and offline speed of information diffusion, at least more so than other online social networks.

The emergence of assortative mixing patterns from high-degree individuals in online social networks aligns with social network analysis theories that predate the advent of digitized human connection. From a theoretical standpoint, the notion that a small number of people can influence the topology of a network is well-established (Barabási & Bonabeau, 2003). Simply put, popular individuals can attract connections and springboard information spread with their vast network, whereas newer, less popular users have fewer connections and therefore fewer opportunities to spread information.

Furthermore, the propensity for people to preferentially attach is an additional factor of how degree assortativity emerges in a social network. In general, people tend to form connections with well-connected people, thus leading to a rich getting richer effect (Barabási & Albert, 1999). People attaching preferentially often results in human social networks featuring a small number of highly connected nodes, while most nodes have a low degree. Preferential attachment has been found to lead to the emergence of scale-free networks, and because scale-free networks are characterized by power-law degree distributions, they are by definition degree-assortative. Therefore, as preferential attachment drives the emergence of scale-free architecture, by that same token, preferential attachment drives the emergence of degree assortativity.

In Scale-Free Networks, Barabási and Bonabeau (2003) comprehensively distinguish how the presence of just a handful of highly connected nodes can lead to people preferentially attaching, thereby forming degree-assortative networks that mature into scale-free networks. In scale-free networks, the few nodes with the highest degrees tend to connect with each other, and the formation of a core of high-degree nodes can significantly impact degree assortativity, the topology of a network, and the behavior of a network writ large (Barabási & Bonabeau, 2003).

4.2.2 Offline versus online social networks

Numerous studies suggest that offline social networks tend to have higher assortativities than online social networks (Hu & Wang, 2009), and there appears to be one central theory that strives to explain this distinction. G. Wang et al., 2014 suggest that when people move between our online versus offline social networks, they are not changing any individual properties about themselves, but rather entering communities that are built around forming connections in a distinctly online way. "By formalizing our offline social relationships into digital form, [online social networks] have greatly expanded our capacity for social interactions, both in volume and frequency" (G. Wang et al., 2014). Simply put, people form connections differently online than they do offline, which leads to disparities in measures of large-scale network structure, such as network density and degree assortativity (Jaakonmäki et al., 2017; Kwak et al., 2010; G. Wang et al., 2014). In context, the digitization of human connections in online social networks is not the same across online social networks because they have different assortativities. This simulation experiment aims to uncover what effect, if any, this assortative disparity has on how quickly information can spread through online circles.

4.2.3 Variations in assortativity between online social networks

While online social networks tend to exhibit disassortative or neutral mixing by degree G. Wang et al., 2014, studies show homophilic contents spread online, such as users' happiness scores (Bliss et al., 2012) and their number of suicide-related verbalizations (Cero & Witte, 2020) show assortative mixing patterns. Further, far from all popular online social networks are degree-disassortative – for example, sites like Facebook, MySpace, and Flickr are degree-assortative (Mutlu & Garibay, 2020). In addition, a popular Chinese networking platform, Wealink, displayed a shift from degree-assortativity to degree-disassortativity between 2005 and 2007 (Hu & Wang, 2009). Thus, not only does the assortativity of online social networks differ from that of offline social networks, but online social networks exhibit a range of assortativity. Therefore, degree assortativity in online social networks has the propensity to change.

Knowing how degree assortativity affects the speed of information diffusion could lead to online social network companies adjusting how people form connections in order to drive faster dissemination of information.

4.3 Information diffusion in online social networks

In addition, online social networks are particularly effective at spreading information – "online social networks play a major role in the spread of information at very large scale" (Guille et al., 2013; Kwak et al., 2010; Mutlu & Garibay, 2020). Popular online social networks – which spread information – have been observed to be assortative, neutral, and disassortative. While this could imply that assortativity plays no role in the spread of information, this is highly unlikely to be the case because a substantial body of literature supports the notion of assortativity as a major influencer in processes of contagion (i.e., information diffusion) across social networks (Bliss et al., 2012; Mutlu & Garibay, 2020). Generally, while information can diffuse no matter the assortativity of a network, evidence suggests that assortativity strongly influences the fraction of the network information can reach.

Figure 2 is a visualization of the relationship between the degree distribution of a network and the size of its giant component (the largest connected cluster) across varying levels of assortativity. This figure suggests that disassortative online social circles might spread information faster as they tend to create larger connected components at higher mean degrees, which is what tight-knit online social circles look like (Mutlu & Garibay, 2020).

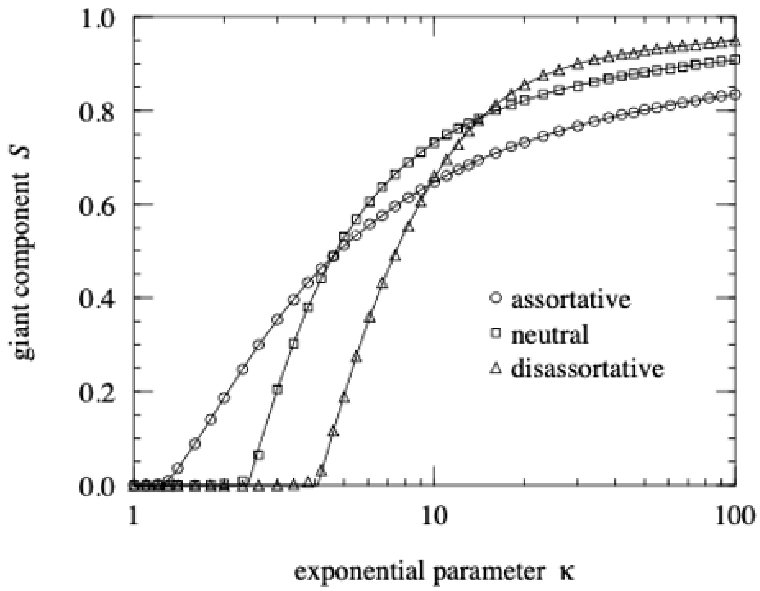


Figure 2: Percent of nodes in the giant component, S, versus the mean degree of the network, κ (Newman & Girvan, 2002)

4.3.1 Does assortativity affect the speed of information diffusion?

In principle, for networks in which the average person only has a few connections, the network tends to be more interconnected when it is assortatively mixed by degree (because the giant component is larger). Moreover, as the average node grows in degree, disassortative networks yield the largest giant component compared to assortatively and neutrally mixed networks (Newman & Girvan, 2002). This mechanism has far-reaching implications for processes of contagion, such as information diffusion, due to the presence of core-periphery architecture in assortative networks (Barabási & Bonabeau, 2003; Newman & Girvan, 2002).

Assortative networks exhibit a core-periphery anatomy across all levels of network density (i.e., the average degree of a node) (Newman & Girvan, 2002) – the popular nodes that connect with other popular nodes are the core, and the unpopular nodes are the periphery. While this can allow for interconnectedness among the core at low levels of network density, this also makes connections with the periphery difficult, even at high levels of network density.

> In epidemiological terms, ... assortative networks will support the spread and persistence of a disease more easily than disassortative ones, because they possess a core group of connected high-degree vertices. But the disease is also restricted mostly to that core group. In a disassortative network, although percolation and hence epidemic disease requires a denser network to begin with, when it does happen it will affect a larger fraction of the network, because it is not restricted to a core group(Newman & Girvan, 2002).

Extrapolating disease contagion to the spread of information, information may quickly spread throughout an assortatively mixed network's core but could encounter difficulties in spreading to the periphery of the network. On another hand, because disassortative networks do not have core-periphery architecture, information may spread to a larger portion of a social network, but this demands high network density.

Because online social networks tend to be disassortative and online connections can be formed faster than offline connections, evidence suggests that disassortative networks are perhaps better suited for diffusing information throughout the entirety of an online social network (Mutlu & Garibay, 2020; G. Wang et al., 2014; T. Wang et al., 2018).

Though differences in assortativity among online social networks are well-known, the specific structural and functional role that assortativity plays in information diffusion in online social networks is difficult to investigate given the vast contextual differences between online communities. In part, differences in functionality between online social networks – such as anonymity, user engagement, content posting, interaction, and temporal features (G. Wang et al., 2014) – point to an experiment as the only viable method of investigating the relationship between assortativity and information diffusion in online social networks.

5 Data and Methods

There are several ways to define how those in online social networks are connected. Though Facebook obviously is based around friends, people can form connections through liking the same content, commenting on the same post, for instance. There is an arms race against social network analysis practicians versus private online social network companies for methods of finding and defining friendships.

5.1 Data

Collecting data on who is friends with whom in online social networks is particularly difficult without access to private company data. However, in some rare circumstances, scientists have been allowed to record the friendships of a small handful of willing participants.

This thesis uses Facebook friend network data that was collected by Mcauley and Leskovec (2014) and was made publicly available through the Stanford Network Analysis Project Leskovec and Krevl (2014). In "Discovering social circles in ego networks," Mcauley and Leskovec developed a Facebook application that allowed them to collect the friend network data of ten Stanford graduate students. This application yielded ten distinct circles of Facebook friends, which are depicted in the following plot and described in the table thereafter.

Table 2: Descriptive network statistics of the Facebook social circles graph

Statistic	Value
Number of Nodes	4039
Number of Edges	88234
Density	0.011
Average Degree	43.691
Diameter	8
Clustering Coefficient	0.519
Assortativity	0.064
Average Path Length	3.693
Degree Scaling Exponent	2.51
Number of Connected Components	1

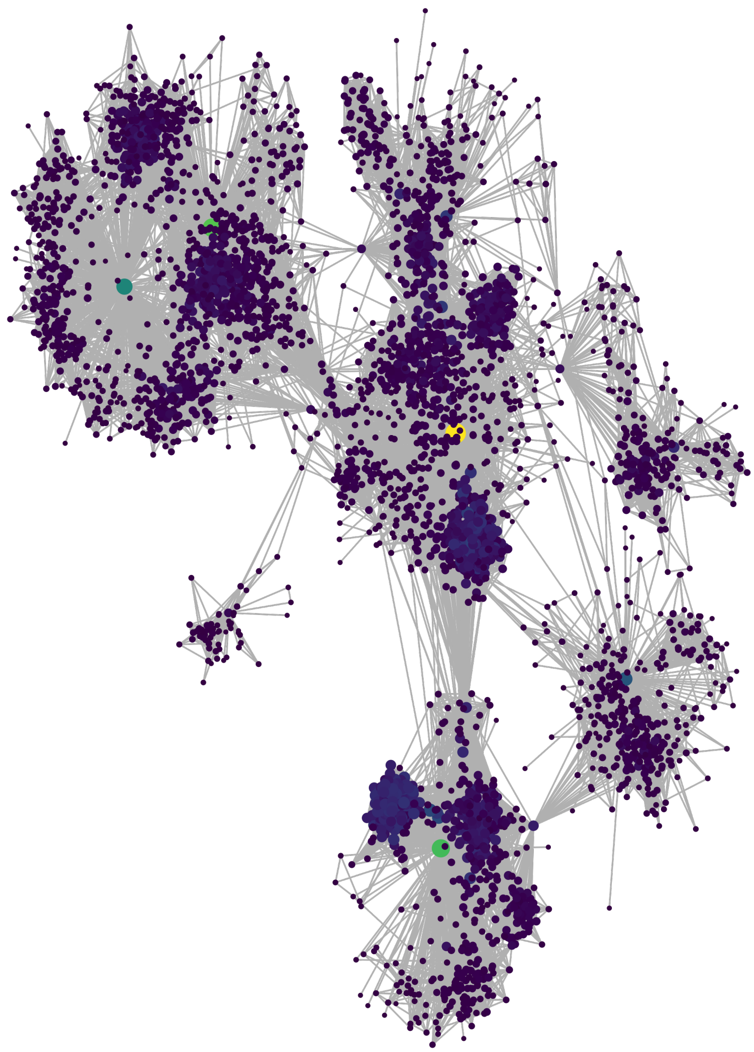


Figure 3: Plot of the Facebook social circles graph. Higher degrees are represented with lighter colors.

Though part of one connected component, the ten circles of Facebook friends appear somewhat separated from one another. Because this network does not include the important friendships that might connect one circle of friends to another, broadly speaking, this collection of ten loosely connected circles of Facebook friends does not represent how, in the anatomy of the Facebook network, neighborhoods of users are surprisingly dense (Ugander et al., 2011).

To make this valuable Facebook network data appropriate, the smallest circle of friends out of the ten in the data set was isolated. Simulating the process of information diffusion through a small group of Facebook friends offers several methodological advantages, one such advantage being that isolating the smallest circle can more truthfully inform how assortativity influences the speed of information diffusion within close-knit circles of friends. Because the mechanisms of diffusion through online social networks are very much an active body of research, the ways in which information diffuses in online social networks become more elusive as it scales.

In addition, selecting just one Facebook circle also significantly limits the computational resources required to rewire the network and simulate processes of information diffusion. Designing a simulation experiment that is not computationally exhaustive not only adds convenience to this experimental methodology, but also repeatability. A plot of the isolated Facebook social circle subgraph is displayed below, followed by descriptive network statistics.

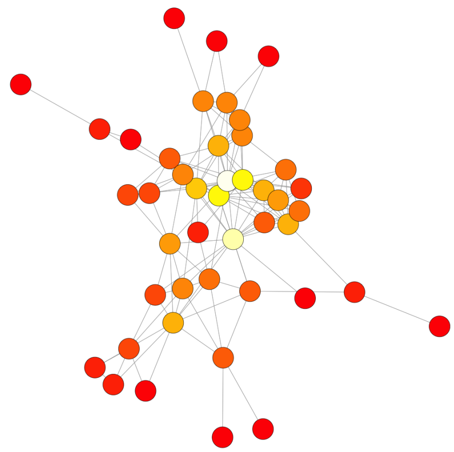


Figure 4: Plot of the Facebook social circles subgraph. Higher degrees are represented with lighter colors.

Table 3: Descriptive network statistics of the Facebook social circles subgraph

Statistic	Value
Number of Nodes	44
Number of Edges	276
Density	0.292
Average Degree	12.545
Diameter	5
Clustering Coefficient	0.444
Assortativity	0.053
Average Path Length	2.567
Degree Scaling Exponent	4.558
Number of Connected Components	1

This subgraph features an assortativity relatively similar to the assortativity of the entire Facebook network ($r = 0.116$) as measured by G. Wang et al., 2014. Therefore, this subgraph is more relevant to the assortative focus of this project while also being more computationally manageable.

5.2 Methods

5.2.1 Measures

The independent variable of this experiment will be the coefficient of degree assortativity as measured by the Pearson correlation coefficient. Observing the fraction of nodes who know the information being diffused after each round will allow the dependent variable – speed of information diffusion – to be calculated.

5.2.2 Rewiring Procedure

Changing the assortativity of a network while minimizing change in other attributes in network shape is complicated. One approach to empirically manipulating assortativity is through random rewiring while holding the degree distribution constant. This is possible by employing the *rewire*() and *keeping_degseq*() functions in igraph (Csardi et al., n.d.). The degree distribution of a network is a key influencer of its shape, so holding it constant throughout a randomized rewiring process helps minimize the change in a network's architecture as it relates to the number of connections each node has (Shirado et al., 2019). Below is the degree distribution of the Facebook subgraph.

This degree distribution indicates that the vast majority of Facebook users in

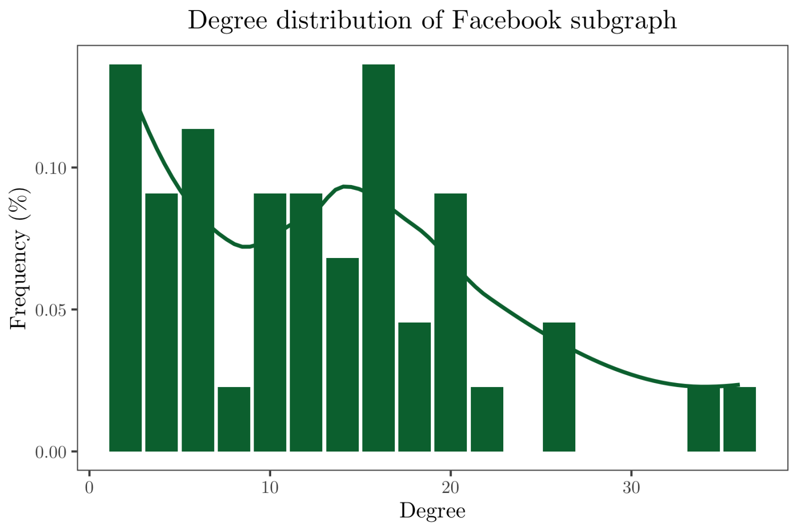


Figure 5: Degree distribution of the Facebook social circles subgraph

this circle have between zero and twenty friends while a few users have thirty-five or more Facebook friends. Such a distribution – in which most nodes have a handful of connections and a handful of nodes have most connections – is indicative of scale-free architecture. When a network is scale-free, it means that there is a very small proportion of nodes with such high degrees that they exceed the scale of the degree distribution of the other nodes (Barabási and Bonabeau, 2003).

Both scale-free architecture and degree assortativity are calculated based on the degrees of the nodes in a network – therefore, minimizing changes in the degree distribution of the Facebook subgraph is imperative to the success of this research. How can one preserve this foundational aspect of the architecture of the Facebook subgraph while also experimentally manipulating assortativity?

As aforementioned, Csardi et al., n.d. authored a rewiring algorithm that allows one to randomly rewire the connections in a network to experimentally manipulate degree assortativity while maintaining the degree distribution of the network. The process of rewiring every connection in the Facebook subgraph was conducted 10,000 times and the assortativity was recorded after every round of rewiring. Below is a histogram of the range of assortativity achieved through randomized rewiring while holding the degree distribution constant.

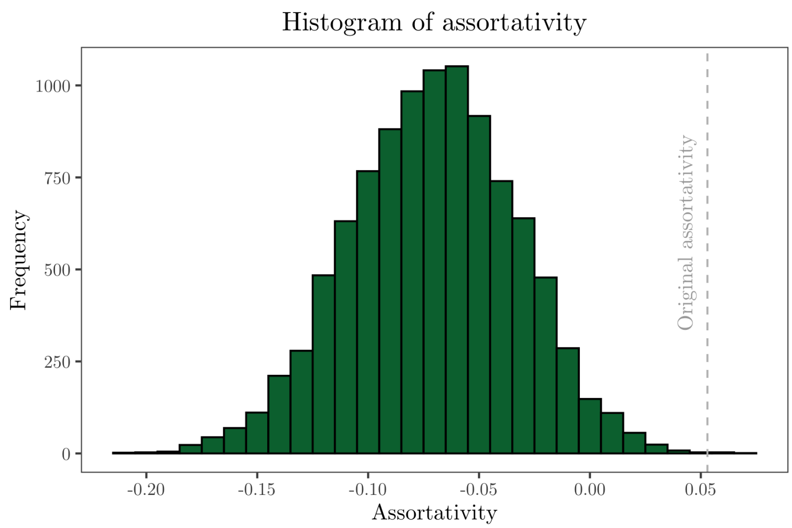


Figure 6: Histogram of assortativity after 10,000 trials

This distribution of assortativity is centered around a mean of $r = -0.03$, far from the observed assortativity of $r = 0.05$ of the Facebook subgraph. A difference of means test, the results of which are presented in the subsequent section, can test whether the Facebook subgraph has a significantly higher assortativity than would be expected through random rewiring.

For the purposes of this simulation experiment, this range of assortativity is unsatisfactory, as it fails to include assortativities greater than 0.1, which have been observed in online social networks such as Facebook (Ugander et al., 2011). Simply put, obtaining assortativity values beyond the bounds achieved with 10,000 trials would be extraordinarily unlikely.

Thankfully, the ability to further change degree assortativity as an independent variable without changing the overall degree distribution is also feasible through incorporating two rewiring algorithms pioneered by Xulvi-Brunet and Sokolov in 2005. In theory, these algorithms allow for changing the assortativity of a graph anywhere between -1 and 1 while holding the degree distribution constant. The algorithms are explained below:

> Starting from a given network, two links of the network connecting four different nodes are randomly chosen at each step. We consider the four nodes associated with these two links, and order them with respect to their degrees. Then, with probability p, the links are rewired in such a way that one link connects the two nodes with the smaller degrees and the other connects the two nodes with the larger degrees; otherwise the links are randomly rewired. In the case that one or both of these new links already existed in the network, the step is discarded and a new pair of edges is selected. This restriction prevents the appearance of multiple edges connecting the same pair of nodes. A repeated application of the rewiring step leads to an assortative version of the original network. Note that the algorithm does not change the degree of the nodes involved and thus the overall degree distribution in the network. Changing the parameter p, it is possible to construct networks with different degrees of assortativity. ... A minor change in our algorithm can produce dis-

sortative mixing. As before, we start from a given network and at each step we chose randomly two links of the network. We order the four corresponding nodes with respect to their degrees. Now, however, we rewire with probability p the edges so that one link connects the highest connected node with the node with the lowest degree and the other link connects the two remaining vertices; with probability 1 - p we rewire the links randomly. In case that any of the new links already existed in the network the step is discarded and a new pair of edges selected. Varying the parameter p, it is possible to construct networks with different degrees of dissortativity. As before, the procedure does not change the degree distribution of the network and does not lead to the appearance of multiple and self-connections" (Xulvi-Brunet & Sokolov, 2005).

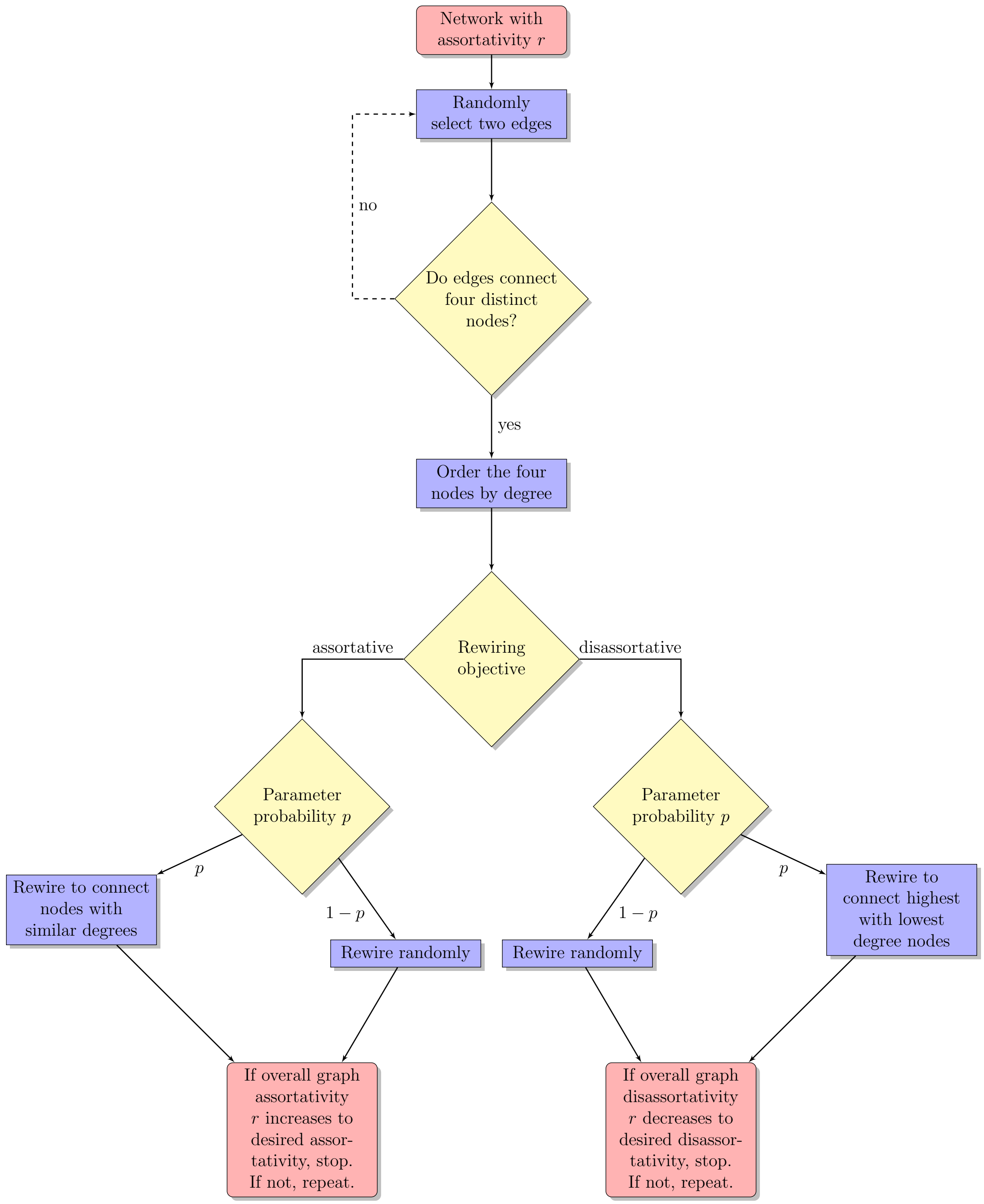
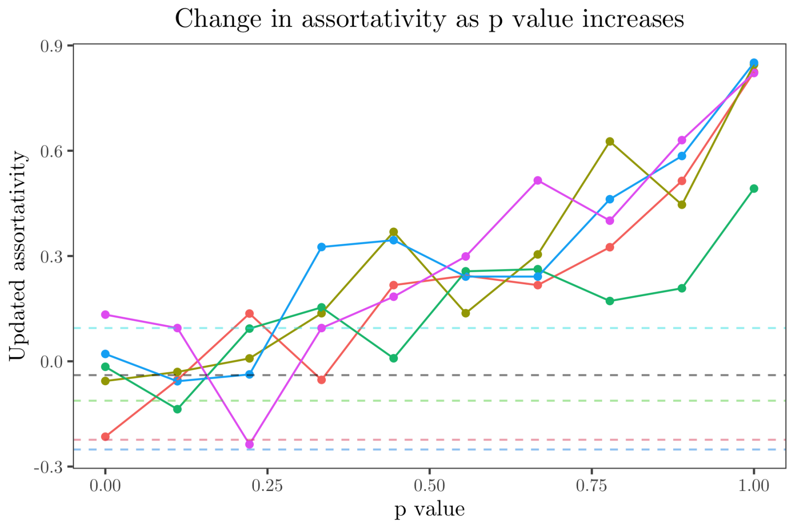
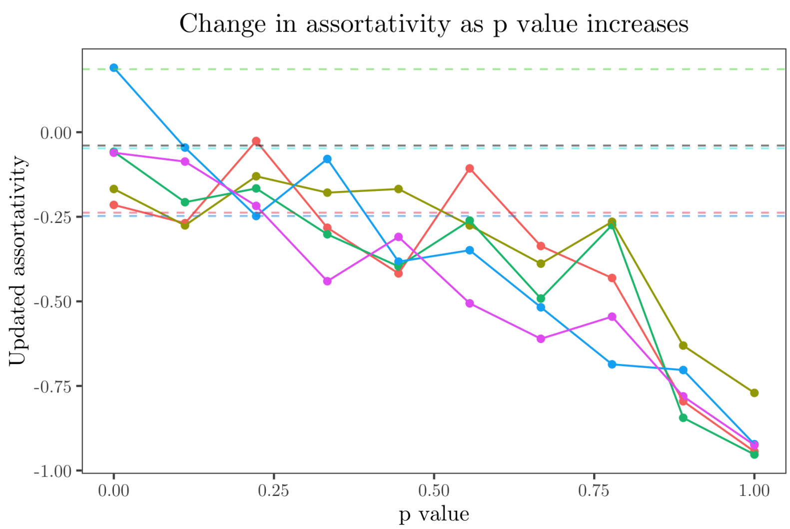


Figure 7: Flowchart of assortative and disassortative rewiring algorithms (Xulvi-Brunet & Sokolov, 2005)

Using R, I then applied these two algorithms to the Facebook subgraph to experimentally manipulate degree assortativity while holding the degree distribution constant. To demonstrate that these algorithms did not get lost in translation, five randomly generated networks were rewired each 200 times, randomly deciding p values between 0 and 1. Histograms of the achieved assortativities for both algorithms are depicted below.



(a) Assortative algorithm



(b) Disassortative algorithm

Figure 8: Scatterplots of assortativity after 100 trials of Xulvi-Brunet and Sokolov rewiring algorithm on 5 Erdos-Renyi random graphs with 30 nodes and a 10% probability of drawing an edge between two nodes. Dotted lines inidcate original assortativities.

One unforeseen idiosyncracy of the Xulvi-Brunet and Sokolov algorithms is that they seem to be less effective at rewiring dense, tight-knit networks to strong levels of either assortativity or disassortativity. After 50 trials of the assortative model of the Xulvi-Brunet and Sokolov algorithms – this time on the Facebook subgraph – a significantly smaller range of assortativities were achieved than were achieved with the sparser Erdos-Renyi random graphs.

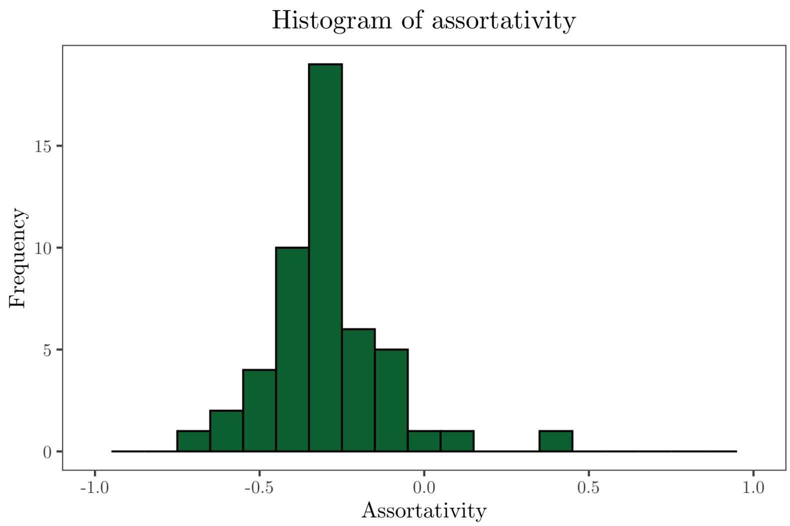


Figure 9: Histogram of assortativity after 50 trials of Xulvi-Brunet and Sokolov rewiring algorithm on Facebook social circle subgraph

Perhaps this phenomenon occurs because the algorithm breaks a rewiring step if any two edges do not connect four distinct nodes, and because the Facebook subgraph is so dense it is breaking the majority of rewiring steps, thus increasing the proportion of randomly rewired edges in relation to assortatively or disassortatively rewired edges. When running the rewiring algorithm on a sparser graph (such as Figure 6, in which choosing four unconnected nodes is more likely), the Xulvi-Brunet and Sokolov algorithm produces a range of assortativities that are impossible via random rewiring.

Despite the Xulvi-Brunet and Sokolov rewiring procedures encountering difficulty in obtaining the highest levels of assortativity, it still expanded the range of assortativity that was possible through the randomized rewiring procedure after 160 trials of rewiring the Facebook subgraph. Figure 10 presents a scatterplot of all 160 rewired graphs, the p values used to rewire them, and the assortativity of the Facebook subgraph.

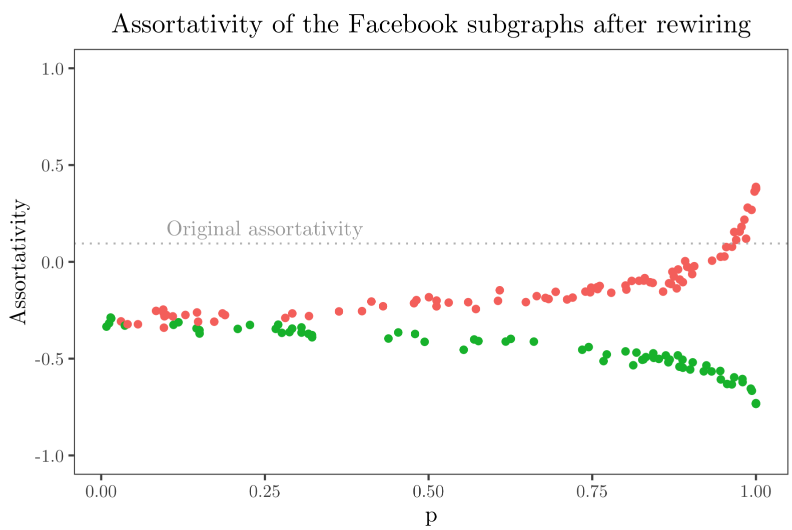
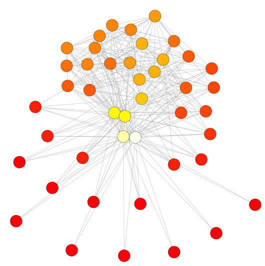
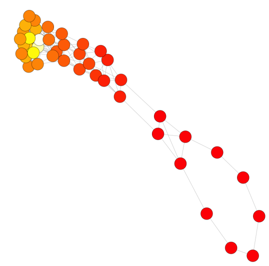


Figure 10: Histogram of assortativity after 160 trials of Xulvi-Brunet and Sokolov rewiring algorithm on Facebook social circle subgraph

Perhaps the most intriguing artifact of this data visualization is the divergence of assortativities as p increases. These two trends represent the two rewiring algorithms. Across the 160 trials, the assortative or disassortative rewiring algorithm was chosen at random, and as demonstrated in Figure 8, higher p values increase the change in assortativity for both the assortative and disassortative algorithms. The most assortative and disassortative rewired Facebook subgraphs are plotted in Figure 11, and their respective assortativities represent the range of assortativity achieved with the Xulvi-Brunei and Sokolov rewiring algorithms.



(a) $r = -0.733$



(b) $r = 0.387$

Figure 11: Minimal and maximal assortativities

Thus, the Xulvi-Brunet and Sokolov rewiring procedure proved effective at experimentally manipulating the assortativity of the Facebook graph while maintaining a consistent degree distribution. However, controlling the degree distribution does not imply that other aspects of the structure of a network change with assortativity. Specifically, average path length, clustering coefficient, shell structure, and percolation properties change significantly with changes in assortativity (Xulvi-Brunet & Sokolov, 2005). Average path length and clustering coefficient are two relevant confounders that are addressed by controlling for them in a regression in the subsequent section.

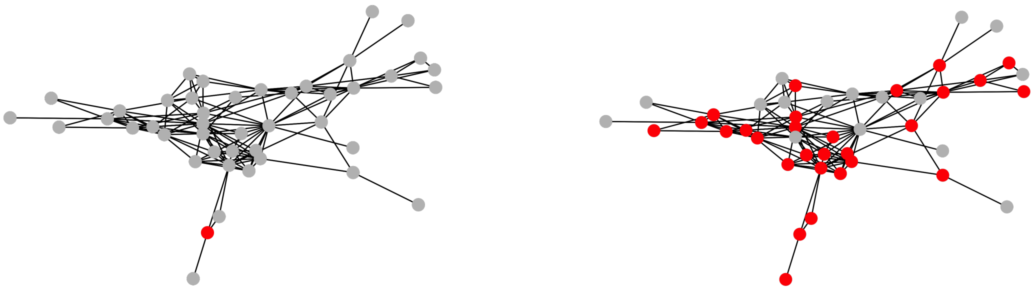
5.2.3 Threshold model of information diffusion

Spreading information is a process of diffusion through a network. As aforementioned, the ways in which things spread through a network is an active area of research. The simplest model of information diffusion is a threshold model, which suggests that people adopt ideas after a certain proportion of their immediate friends have adopted it.

According to Professor Hirokazu Shirado, simpler simulated models of information diffusion through a network allow for a greater extent of reproducability and compatibility with several schools of thought. Countless models of information diffusion encompass the field of social network analysis, and as one creates a more sophisticated model of information diffusion, it also polarizes that model from areas of research that might not adhere to that school of thought.

Thus, the diffusion model adopted is not far from a simple flip of a coin. In the first round of information diffusion, a node is selected at random to be the first person to be 'infected' with the information. For every subsequent round of infection, the infected node spreads the information to each of their connections with a ten percent chance, independently.

Diffusing information is a process of infection (G. Wang et al., 2014; T. Wang et al., 2018). For each iteration of information diffusion, a node was selected at random to be the first and only infected node in the graph. Then, each of the seed node's connections were infected with a ten percent probability. Below is a visual demonstration of the process of infection spread throughout the original Facebook subgraph.



(a) Round 2/12 (b) Round 8/12

Figure 12: Process of infection spread

Rounds of infection are continued until at least 80% of the network becomes infected – this choice of 80% is slightly arbitrary, but takes into account that those last 20% of nodes are not relevant to the process that led to the majority of the network becoming infected.

In summary, using the Xulvi-Brunet and Sokolov rewiring algorithms from 2005, 160 versions of the Facebook subgraph were created that range in assortativity from $r = -0.733$ to $r = 0.387$. These networks all have the save degree distribution. Information will be diffused information via a threshold diffusion model, which offers simplicity and repeatability.

6 Results

Broadly, the results of this thesis support the hypothesis that the speed of information diffusion and assortativity have an inverse relationship. That is, as assortativity increases, the speed of information diffusion decreases. In addition, results indicate that friends in the Facebook subgraph form connections that increase degree assortativity statistically significantly higher than the same number of random connections would. Moreover, results suggest that assortativity predicts the speed of information diffusion with strong statistical significance, whereas other measures of large-scale network structure that correlate with assortativity, such as average path length and clustering coefficient, do not.

Table 4: Facebook subgraph assortativity versus randomly rewired assortativity

t value	df	p value	Confidence Interval	
-3.291	9999	0.001	-0.145	0.003

$\alpha = 0.05$

The difference of means test summarized in Table 4 suggests that Facebook users in the social circle form connections assortatively with strong statistical significance. Indeed, Facebook users in the social circle tend to form connections with popular, highly connected users, and this phenomenon is extremely unlikely to be due to random chance. Does this elevated assortativity of the Facebook subgraph offer an advantageous speed of information diffusion? Simulation results, discussed hereafter, suggest that the speed of information diffusion of the original Facebook subgraph is relatively slow.

To investigate how assortativity affects information diffusion, the threshold diffusion process was repeated 100 times for each of the 161 subgraphs (1 original, 160 rewired with the Xulvi-Brunet and Sokolov rewiring algorithms). The density curve below depicts how many rounds, on average, it took for each class of rewired graph to reach 80% diffusion. Here, "class" implies that when the subgraph is rewired to have a greater assortativity, it is classified as assortative, and vice versa.

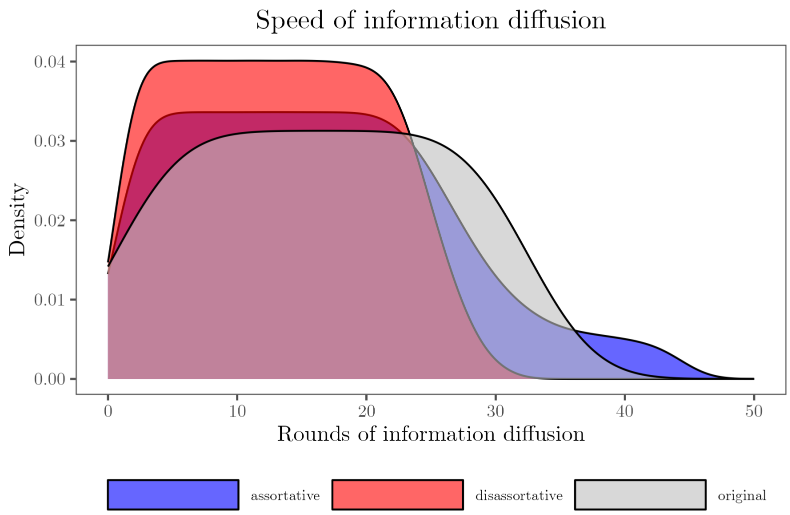


Figure 13: Density curve of how many rounds it took to reach 80% infection for each class of network

Regardless of direction or magnitude, when rewired, the Facebook subgraphs tended to diffuse information more quickly than the original version. This finding suggests that the assortativity of the Facebook subgraph, $r = 0.05$, lends itself to a process of information diffusion that could be accelerated with assortative or disassortative rewiring.

The fact that the disassortative class of rewired graphs has a greater density between rounds 0 through 100 indicates that, on average, when the Facebook graph was rewired disassortatively, regardless of magnitude, it tended to spread information faster than assortatively rewired graphs. Moreover, on average, and under a threshold model of diffusion, information diffuses fastest when a Facebook subgraph is rewired disassortatively.

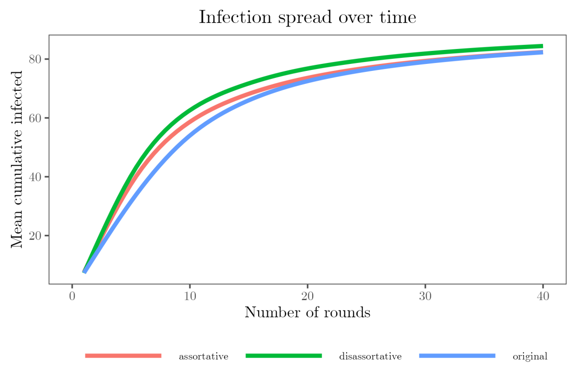


Figure 14: Smoothed regression of cumulative percent infected versus number of rounds for each of the networks

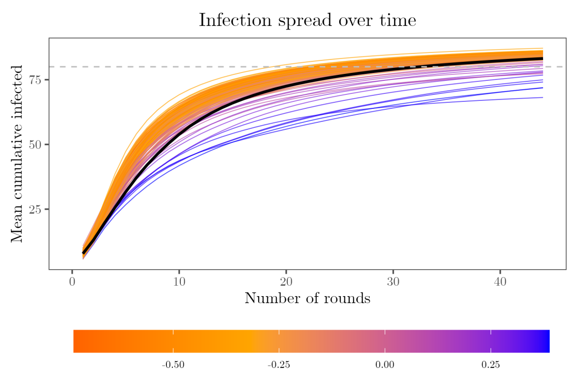


Figure 15: Scatterplots of cumulative percent infected versus number of rounds for each of the networks

On average, disassortatively rewired subgraphs diffused information faster than assortatively rewired subgraphs. Moreover, the original Facebook subgraph diffused information slower than the average disassortatively or assortatively rewired subgraph.

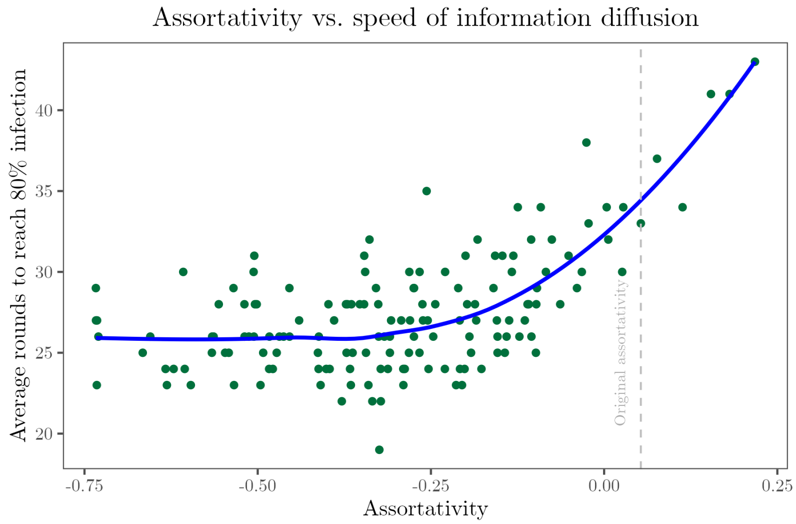


Figure 16: Assortativity versus speed of information diffusion

In Figure 16, the average number of rounds to reach 80% infection for each of the 161 subgraphs was plotted against their respective assortativities, and a loess regression was fitted to the direct correlation. As shown, assortativity appears to have a profound impact on the number of rounds it takes to reach 80% infection. Moreover, variations in arbitrary parameters of the process of infection spread appear to have no substantive effect on the relationship between assortativity and the speed of information diffusion.

The first arbitrary decision in the infection spread process is the probability of a node infecting a given connection – at the selected probability of ten percent, the infection process is slow enough to illustrate a more pronounced difference in the speed of information diffusion by the degree assortativity of each subgraph.

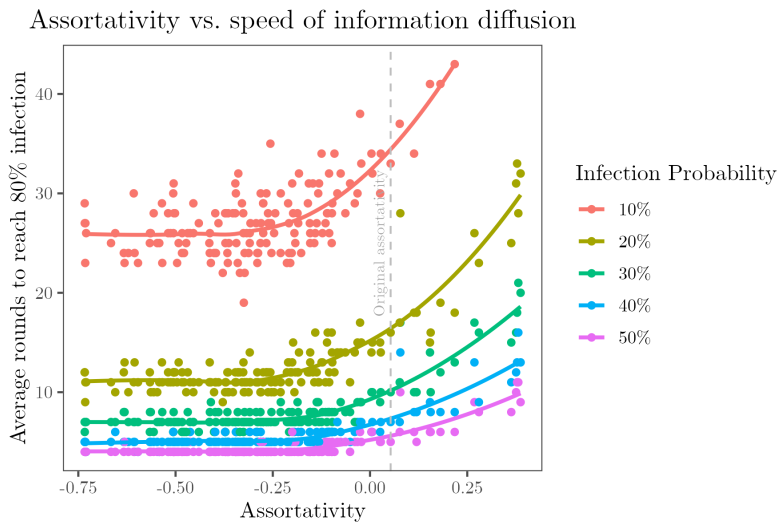


Figure 17: Assortativity versus speed of information diffusion for varying infection probabilities

At faster infection rates, information still passes faster in disassortative networks, and the overall shape of the relationship between assortativity and the speed of information diffusion remains consistent. At some point, however, the probability of infecting a connection can become so high that differences in assortativity have a diminishing effect on the speed limit of information diffusion.

Secondly, that the simulated infection process randomly selects a node to begin the infection process brings into question how a different method of selecting the first infected node might change the observed strong, inverse relationship between degree assortativity and the speed of information diffusion. An alternative selection method is iterating through an ordered list of each node to guarantee that every node is tested – such a method does not yield a substantive change in assortativity and information diffusion.

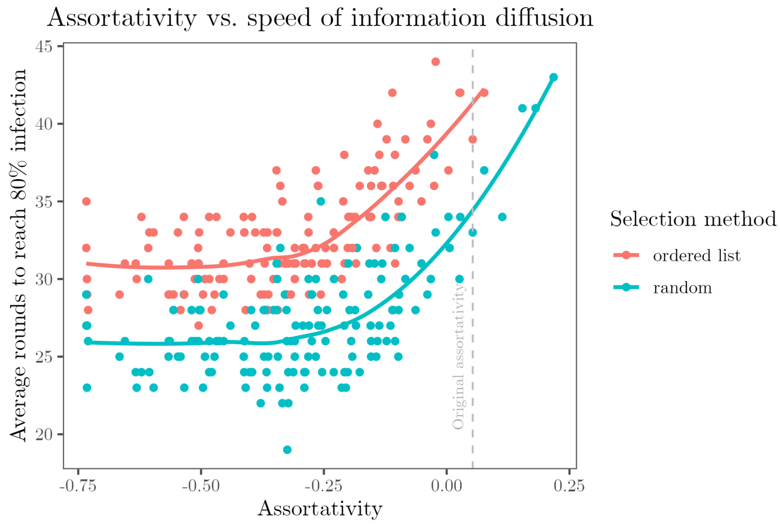


Figure 18: Assortativity versus speed of information diffusion for varying selection methods

When selecting the first infected node from an ordered list versus random selection, a similar phenomenon as illustrated in Figure 17 occurs, in which the strong, inverse relationship between assortativity and the speed of information is preserved with a minor change in the y-intercept. The difference in rounds between the ordered list selection method and the random selection method is consistently approximately 5 rounds across the distribution of assortativities, which suggests that either selection method would yield a similar correlation coefficient for degree assortativity versus average rounds of diffusion.

Third, the relationship between assortativity and the speed of information diffusion does not appear to be determined by the exact subgraph selected. The simulation experiment was repeated for a different Facebook subgraph of similar size and degree assortativity, but of a greater density, lesser average path length, and greater clustering coefficient.

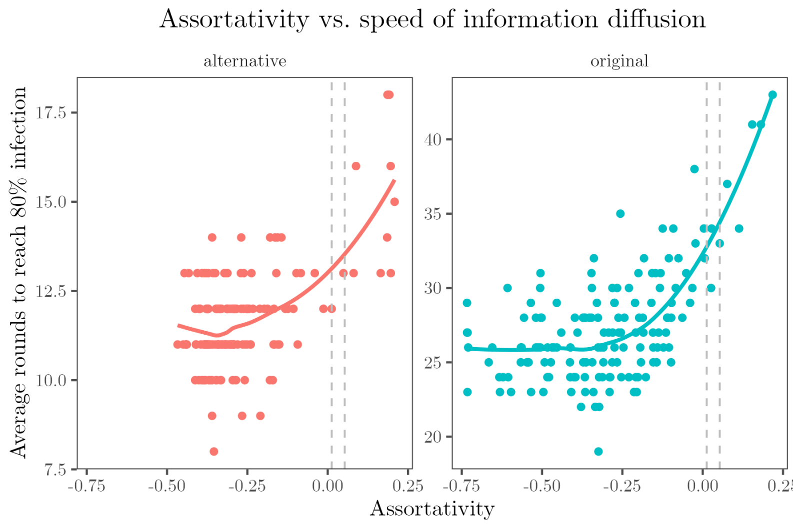


Figure 19: Assortativity versus speed of information diffusion for varying network choices

Again, the overall trend is preserved, but as expected, due to the density of the alternative subgraph being nearly double that of the original subgraph, the Xulvi-Brunet and Sokolov, 2005 rewiring algorithm yielded a more limited set of rewired assortativities. In addition, because the greater density fosters more chances for infection to spread, the overall speed of information diffusion for the alternative network is expectedly faster.

Lastly, the threshold of information diffusion, set at 80%, governs when the simulated information diffusion process ends.

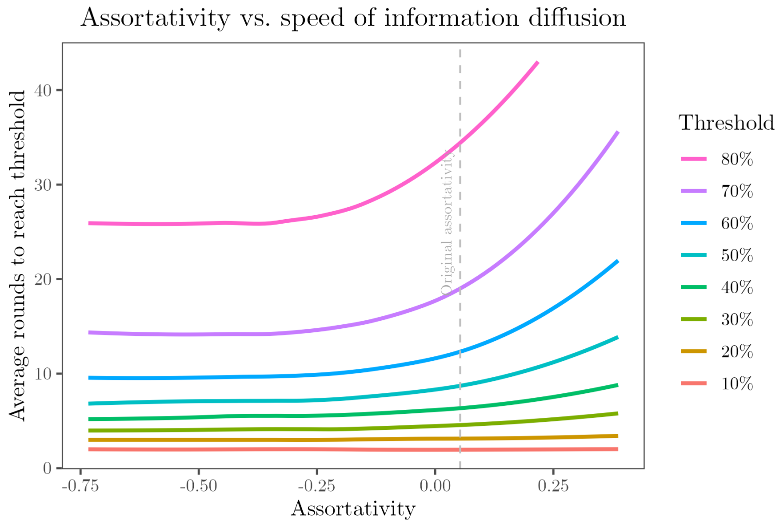


Figure 20: Assortativity versus speed of information diffusion for varying saturation thresholds

Again, similar to Figure 17, the overall trend between assortativity and the speed of information diffusion is generally preserved despite changes in this parameter of infection spread. At the lowest thresholds, the simulated infection process stops too early for the topological diffusion mechanisms between degree-assortative and degree-disassortative graphs to emerge.

In sum, the ten percent chance of infection, random starting node selection, eighty percent threshold, and specific network choice are in place largely for illustrative purposes – when these four controls are varied, results are substantively similar. At some of the more extreme values, the infection process occurs too quickly for a visible distinction between assortativity and speed of information diffusion to emerge.

Moreover, a linear regression of the same relationship was performed, indicating a strong positive correlation between assortativity and rounds to reach 80%

diffusion at all standard levels of significance. The first regression model indicates that, on average, a 10% increase in assortativity is correlated with an increase in the number of rounds of information diffusion of 0.67.

Furthermore, a standardized regression was performed to test the statistical significance of the relationship between assortativity and the speed of information diffusion while also controlling for average path length and clustering coefficient, which Xulvi-Brunet and Sokolov found to correlate with changes in degree assortativity (2005). The process of standardizing each variable in the regression involved dividing each variable by its standard error.

Table 5: Regression Results

	Model 1: Linear regression		Model 2: Standardized regression	
	Coef.	P-value	Coef.	P-value
Intercept	19.13	< 0.001***	-0.01	0.93
Updated Assortativity	6.71	< 0.001***	0.46	< 0.001***
Avg Path Length			0.16	0.12
Clustering Coefficient			-0.12	0.08
R-squared	0.32		0.35	
Adj. R-squared	0.32		0.33	
F-statistic	73.62		27.46	
P-value (F-stat)	< 0.001***		< 0.001***	

$\alpha = 0.01$

Notice that the average path length and the clustering coefficient do not have a statistically significant effect on the speed of information diffusion in model 2. Such a result indicates that, while a friend circle's assortativity might change with its average path length and clustering coefficient, these measures of the shape of a social network are unlikely to have as much as an effect on the speed of information diffusion as degree assortativity does.

The intercept and correlation coefficient of linear regression model 1, in addition to supplemental coefficients based on varying infection probabilities, is extrapolated to predict how fast information might diffuse in small, tight-knit circles of users in online social networks beyond Facebook based on differences in their assortativities. The predicted speed of information diffusion for Twitter, YouTube, Instagram, Reddit, and LinkedIn are depicted in Figure 21.

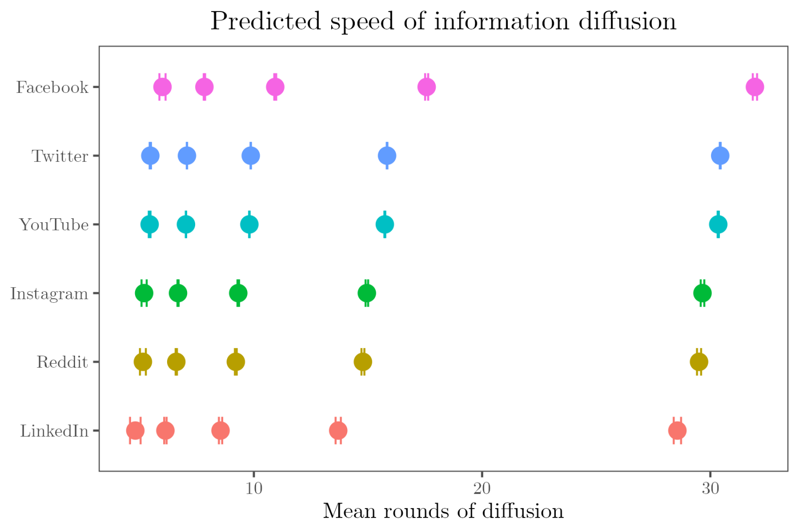


Figure 21: Predicted speed of information diffusion for popular online social networks across varying infection probabilities

LinkedIn is the least assortative network compared to the other five – given its disassortative architecture, perhaps circles of LinkedIn users could spread information roughly ten percent faster than a similar circle of users on Facebook, for instance.

Although a ten percent difference in the speed of information diffusion might appear underwhelming, at scale, small increases in the speed of information diffusion can make a powerful impact.

7 Conclusions

The objective of this thesis is to investigate what, if any, influence degree assortativity has on the speed of information diffusion in online social networks. In summary, the past twenty years of research on degree assortativity indicates that different popular online social networks feature different degree assortativities and that these differences in assortativity might have an effect on how ideas diffuse through a network.

This project aims to demonstrate, with a simulation experiment, that degree assortativity has a strong effect on the speed of information diffusion through tight-knit social circles. To create this simulation experiment, I rewired a 44-node Facebook social circle using two rewiring algorithms that allowed me to change assortativity while maintaining the same degree distribution. This process yielded 160 alternative versions of the original graph and, in essence, I observed how fast each of them could diffuse the same information.

The results of my simulation experiment suggest that the more degree-disassortative a network is, the faster it can spread information. Therefore, because online social networks differ in degree assortativity, it is likely that the same information among the same group of users could diffuse faster on a more disassortative platform. In summary, evidence suggests different social media platforms likely diffuse information at different speeds which are governed by their assortativities.

Investigating how assortativity influences the speed of information diffusion in online social networks contributes to a growing body of literature that seeks to connect the structure and function of online social networks to the diffusion of information. This has far-reaching implications for addressing how behaviors (Kim et al., 2015), happiness (Bliss et al., 2012), suicide-related verbalizations (Cero & Witte, 2020), emergency notifications (Lee et al., 2022), and misinformation (Akrouf et al., 2013) spread online.

Online social networks have a profound impact on the friendships we form and the information we consume. As social media platforms incorporate algorithms to drive the content we are exposed to, understanding how small changes in assortativity can significantly impact the spread of ideas could potentially enable them to rewire our social circles to allow us to spread information more quickly with our connections while also preserving the essence of our networks.

8 Appendix

```
#Assortative rewiring algorithm
library(igraph)
assortative_rewire <- function(network, p) {
  g <- network
  num_iterations <- 100 * gsize(g)
  #100 is arbitrary but ensures every node will be selected

  progress_bar <- txtProgressBar(min = 0, max = num_iterations, style = 3)

  for (i in 1:num_iterations) {

    edges <- E(g)
    #Randomly select 2 edges
    edge_indices <- sample(length(edges), 2, replace = FALSE)
    edge1 <- edges[edge_indices[1]]
    edge2 <- edges[edge_indices[2]]
    nodes <- unique(c(ends(g, edge1), ends(g, edge2)))
    #Make a list of the unique nodes connected by the randomly selected edges

```

```
    if (length(nodes) == 4) {
      degrees <- degree(g, nodes)
      sorted_nodes <- nodes[order(degrees)]
      #If the 2 vertices connected by random edge 1 are different from
      #the 2 vertices connected by random edge 2, order them by degree

      if (runif(1) < p) {
        new_edges <- c(sorted_nodes[1], sorted_nodes[2], sorted_nodes[3], sorted_nodes[4])
        #Order nodes by descending degree
      } else {
        #Alternatively, order nodes randomly
        shuffled_nodes <- sample(sorted_nodes)
        new_edges <- c(shuffled_nodes[1], shuffled_nodes[2], shuffled_nodes[3], shuffled_nodes[4])
      }

      new_edge1 <- get.edge.ids(g, c(new_edges[1], new_edges[2]), directed = FALSE, error = FALSE)
      new_edge2 <- get.edge.ids(g, c(new_edges[3], new_edges[4]), directed = FALSE, error = FALSE)

```

```
      if (new_edge1 == 0 && new_edge2 == 0) {
        g <- delete_edges(g, c(edge1, edge2))
        g <- add_edges(g, new_edges)
        #Do any of these new edges already exist? If not,
        #delete the original edges and add either the randomly rewired edges or
        #the assortatively rewired edges, whichever was created above
      }
    }
    setTxtProgressBar(progress_bar, i)
  }
  close(progress_bar)
  return(g)
}
```

```
#Disassortative rewiring algorithm
library(igraph)
disassortative_rewire <- function(network, p) {
  g <- network
  num_iterations <- 100 * gsize(g)
  #100 is arbitrary but ensures every node will be selected

  progress_bar <- txtProgressBar(min = 0, max =
```

```
    num_iterations, style = 3)

  for (i in 1:num_iterations) {
    edges <- E(g)
    #Randomly select 2 edges
    edge_indices <- sample(length(edges), 2, replace =
       FALSE)
    edge1 <- edges[edge_indices[1]]
    edge2 <- edges[edge_indices[2]]
    nodes <- unique(c(ends(g, edge1), ends(g, edge2)))
    #Make a list of the unique nodes connected by the
       randomly selected edges

    if (length(nodes) == 4) {
      degrees <- degree(g, nodes)
      sorted_nodes <- nodes[order(degrees)]
      #If the 2 vertices connected by random edge 1 are
         different from
      #the 2 vertices connected by random edge 2 , order
         them by degree

      if (runif(1) < p) {
        new_edges <- c(sorted_nodes[1], sorted_nodes[4],
           sorted_nodes[2], sorted_nodes[3])
```

```
        #Order nodes alternating from highest to lowest degree
      } else {
        #Alternatively, order nodes randomly
        shuffled_nodes <- sample(sorted_nodes)
        new_edges <- c(shuffled_nodes[1], shuffled_nodes[2], shuffled_nodes[3], shuffled_nodes[4])
      }

      new_edge1 <- get.edge.ids(g, c(new_edges[1], new_edges[2]), directed = FALSE, error = FALSE)
      new_edge2 <- get.edge.ids(g, c(new_edges[3], new_edges[4]), directed = FALSE, error = FALSE)

      if (new_edge1 == 0 && new_edge2 == 0) {
        g <- delete_edges(g, c(edge1, edge2))
        g <- add_edges(g, new_edges)
        #Do any of these new edges already exist? If not,
        #delete the original edges and add either the randomly rewired edges or
        #the disassortatively rewired edges, whichever was created above
      }
    }
```

```
      setTxtProgressBar(progress_bar, i)
    }
    close(progress_bar)
    return(g)
}
```

References

Akrouf, S., Laifa, M., Yahia, B., & Mouhoub, N. (2013). Social network analysis and information propagation: A case study using flickr and youtube networks. *International Journal of Future Computer and Communication*, *Vol. 2,* 246–252. https://doi.org/10.7763/IJFCC.2013.V2.161

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

Barabási, A.-L., & Bonabeau, E. (2003). Scale-free networks. *Scientific American*, *288*(5), 60–69.

Bliss, C. A., Kloumann, I. M., Harris, K. D., Danforth, C. M., & Dodds, P. S. (2012). Twitter reciprocal reply networks exhibit assortativity with respect to happiness [Place: AMSTERDAM Publisher: Elsevier B.V]. *J COMPUT SCI-NETH*, *3*(5), 388–397. https://doi.org/10.1016/j.jocs.2012.05.001

Cero, I., & Witte, T. K. (2020). Assortativity of suicide-related posting on social media [Place: WASHINGTON Publisher: American Psychological Association]. *AM PSYCHOL*, *75*(3), 365–379. https://doi.org/10.1037/amp0000477 Medline

Csardi, G., Nepusz, T., Horvat, S., Traag, V., & Zanini, F. (n.d.). *Igraph r package*. Retrieved June 5, 2022, from https://igraph.org/r/

Ellison, N., Steinfield, C., & Lampe, C. (2007). The benefits of facebook "friends:" social capital and college students' use of online social network sites. *J. Computer-Mediated Communication*, *12*, 1143–1168. https://doi.org/10.1111/j.1083-6101.2007.00367.x

Ferrara, E., Interdonato, R., & Tagarelli, A. (2014). Online popularity and topical interests through the lens of instagram. *HT 2014 - Proceedings of the 25th ACM Conference on Hypertext and Social Media*, 24–34.

Furht, B. (2010). *Handbook of social network technologies and applications* (1st ed. 2010.). Springer US. https://doi.org/10.1007/978-1-4419-7142-5
Includes bibliographical references and index.

Guille, A., Hacid, H., Favre, C., & Zighed, D. A. (2013). Information diffusion in online social networks: A survey [Place: NEW YORK Publisher: Assoc Computing Machinery]. *SIGMOD REC*, *42*(2), 17–28. https://doi.org/10.1145/2503792.2503797

Hu, H.-B., & Wang, X.-F. (2009). Disassortative mixing in online social networks [Place: BRISTOL Publisher: IOP Publishing]. *EPL-EUROPHYS LETT*, *86*(1), 18003–18003. https://doi.org/10.1209/0295-5075/86/18003
publisher-ID:epl11711

Jaakonmäki, R., Müller, O., & Brocke, J. v. (2017). The impact of content, context, and creator on user engagement in social media marketing. https://doi.org/10.24251/HICSS.2017.136

Kim, D. A., Hwong, A. R., Stafford, D., Hughes, D. A., O'Malley, A. J., Fowler, J. H., & Christakis, N. A. (2015). Social network targeting to maximise population behaviour change: A cluster randomised controlled trial. *Lancet*, *386*(9989), 145–153. https://doi.org/10.1016/S0140-6736(15)60095-2

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is twitter, a social network or a news media? *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, *19*. https://doi.org/10.1145/1772690.1772751

Lee, S., Jain, S., & Son, Y.-J. (2022). A hierarchical decision-making framework in social networks for efficient disaster management [Place: NEW YORK Publisher: Assoc Computing Machinery]. *ACM T MODEL COMPUT S*, *32*(1), 1–26. https://doi.org/10.1145/3490027

Leskovec, J., & Krevl, A. (2014, June). *SNAP datasets: Stanford large network dataset collection*. Retrieved March 28, 2023, from https://snap.stanford.edu/data/ego-Facebook.html

Light, R., & Moody, J. W. (2020). *The oxford handbook of social networks*. Oxford University Press

Includes bibliographical references and index.

Marwick, A. (2015). Instafame: Luxury selfies in the attention economy. *Public Culture*, *27*, 137–160. https://doi.org/10.1215/08992363-2798379

Mcauley, J., & Leskovec, J. (2014). Discovering social circles in ego networks [Publisher: ACM]. *ACM transactions on knowledge discovery from data*, *8*(1), 1–28. https://doi.org/10.1145/2556612

Mislove, A., Marcon, M., Gummadi, K., Druschel, P., & Bhattacharjee, B. (2007). Measurement and analysis of online social networks, 29–42. https://doi.org/10.1145/1298306.1298311

Mutlu, E. Ç., & Garibay, I. (2020). Effects of assortativity on consensus formation with heterogeneous agents [Place: Ithaca Publisher: Cornell University Library, arXiv.org]. *arXiv.org*.

Newman, M. (2001). The structure of scientific collaboration networks [Place: WASHINGTON Publisher: National Academy of Sciences]. *P NATL ACAD SCI USA*, *98*(2), 404–409. https://doi.org/10.1073/pnas.021544898

Communicated by Murray Gell-Mann, Santa Fe Institute, Santa Fe, NM

Newman, M., Forrest, S., & Balthrop, J. (2002). Email networks and the spread of computer viruses [Place: United States]. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, *66*(3), 035101/4–035101. https://doi.org/10.1103/PhysRevE.66.035101

Newman, M., & Girvan, M. (2002). Mixing patterns and community structure in networks. In *Statistical mechanics of complex networks* (pp. 66–87). Springer Berlin Heidelberg.

Qian Chen, Hyunseok Chang, Govindan, R., & Jamin, S. (2002). The origin of power laws in internet topologies revisited [Journal Abbreviation: INFCOM]. *INFCOM*, *2*, 608–617 vol.2. https://doi.org/10.1109/INFCOM.2002.1019306

Rajendran, N. (2021). RPubs - LinkedIn Talent Flow Network Analysis. Retrieved May 5, 2023, from https://rpubs.com/nr2462/845876

Shirado, H., Iosifidis, G., & Christakis, N. A. (2019). Assortative mixing and resource inequality enhance collective welfare in sharing networks [Place: WASHINGTON Publisher: National Academy of Sciences]. *P NATL ACAD SCI USA*, *116*(45), 22442–22444. https://doi.org/10.1073/pnas.1911606116

Edited by Mary C. Waters, Harvard University, Cambridge, MA, and approved September 27, 2019 (received for review July 10, 2019)

Simmel, G. (1908). Conflict and the web of group affiliations. *New York, NY: The Free*.

Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). The anatomy of the facebook social graph. *arXiv.org*.

Vega-Oliveros, D. A., da Fontoura Costa, L., & Rodrigues, F. A. (2020). Influence maximization by rumor spreading on correlated networks through community identification. *Communications in nonlinear science numerical simulation*, *83*, 105094–.

Wang, G., Wang, B., Wang, T., Nika, A., Zheng, H., & Zhao, B. (2014). Whispers in the dark: Analysis of an anonymous social network. *Proceedings of the 2014 Conference on internet measurement conference*, 137–150.

Wang, T., Brede, M., Ianni, A., & Mentzakis, E. (2018). Social interactions in online eating disorder communities: A network perspective [Publisher: Public Library of Science]. *PLOS ONE*, *13*(7), 1–17. https://doi.org/10.1371/journal.pone.0200800

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks [Place: LONDON Publisher: MACMILLAN MAGAZINES LTD]. *NATURE*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Xulvi-Brunet, R., & Sokolov, I. (2005). Changing correlations in networks: Assortativity and dissortativity. *Acta Physica Polonica B - ACTA PHYS POL B*, *36*.

Yale Institute for Network Science. (2022, September 29). *YINS seminar: Nicholas christakis 9/21/2022, "social network interventions"*. Retrieved November 19, 2022, from https://www.youtube.com/watch?v=PR21oQpeKZM