From a Hashtag to a Movement:Modeling Conversation Around #MeToo onTwitter

Citation

Istvan, Brooke. 2019. From a Hashtag to a Movement:Modeling Conversation Around #MeToo onTwitter. Bachelor's thesis, Harvard College.

From a Hashtag to a Movement: Modeling Conversation Around #MeToo on Twitter

A Senior Thesis Presented by
Brooke Istvan
To the John A. Paulson School of Engineering and Applied Sciences in partial fulfillment of the requirements for the degree of Bachelor of Arts in the subject of Applied Mathematics

Harvard College
Cambridge, Massachusetts, USA
March 29, 2019

Abstract

In this paper I evaluate the effectiveness of three different model frameworks from epidemiology, SIR, SISR, and SIRS, at capturing the dynamics and the spread of social movements online. For this project I look specifically at data from #MeToo and the corresponding movement. The data suggests that the hashtag movement can be divided into groups of individuals in the U.S. and abroad. I take into account these between group dynamics. I estimate model transition probability parameters from the Twitter data on #MeToo and compare them with optimal parameters found by minimizing mean squared error of the model output with respect to key metrics from the data. I run simulations of the models over an underlying social network that is representative of Twitter and compare results quantitatively and qualitatively to the data. SISR appears to be the most effective single model at capturing the dynamics of both the hashtag and the movement. Extensions are made to consider the effect of exogenous reinjection of infection into the network and network community structure on the model output.

Acknowledgements

This paper benefited from the help and advice of many people, most notably Sarah Iams. Without Sarah as an advisor and without the dynamical systems course she teaches, I would have never undertaken or completed this project. Thank you to Ryan Gallagher, Brooke Foucault Welles, and Chris Danforth for providing me with the sample of #MeToo tweets as well as readings, advice, and support. I would also like to thank the lovely person who broke into the car and stole my backpack with all my first semester thesis work in it over winter break for teaching me the importance of github. Finally, none of this would have been possible if it weren't for the millions of women and men everywhere tweeting #MeToo and bringing matters of sexual assault and harassment into a new public light. Thank you for spreading the infection.

Contents

1 Introduction

Models of infectious spread have long been studied and have many relevant applications. Such models have historically been applied to epidemiology [13, 19, 21, 27] but more and more frequently are being used to study information diffusion and idea spreading in a variety of contexts including political movements, collective actions, fads, innovations, memes, and hashtags [7, 11, 23, 32, 33]. This paper details the construction of extended epidemiological models over a network in an attempt to capture the dynamics of social movements. I construct 3 different models and evaluate their effectiveness at explaining both the spread of #MeToo and the corresponding movement as a whole. I examine Twitter data for #MeToo and use this particular movement as a case study for understanding the spread of influential social movements over a social network.

In the age of social media, more people than ever before are exposed to new ideas and information online. 68% of Americans got at least some of their news on social media in 2018 [20]. The use of social networking sites allows large grassroots movements to spring up and take hold, especially given their popularity with "slacktivists" online [6]. These social movements tend to be of an extremely sensitive nature, and there has been a growing presence of online radicalizing extremism in the age of social media [15]. Twitter in particular has been a platform for political and politicized discourse [9, 14].

Users on Twitter can add hashtags to the end of tweets associated with the content of their tweets and some specific hashtags have reached massive levels of popularity, representing prominent social movements. In America hashtags like #BlackLivesMatter, #TakeAKnee, #TimesUp, and #MeToo have become politicized tools aimed at effecting change in recent years. Similarly, hashtags like #Third_Debate, #ILoveYouChina, #YemenInquiryNow, #BringBackOurGirls and #StopFundingHate have gained popularity in different places around the world corresponding with the increasing influence of the movements they represent [28]. Social media has become so popular that mass media outlets now report on national discourse that takes place online [24]. Major news outlets are also sharing their content on social networking sites themselves [2, 29, 30].

Between foreign influence, fake news, and echo chambers that may emerge in social networks, there is a fear that society is becoming more polarized with the influence of social media, threatening the state of democracy [5, 12]. In this paper I propose and evaluate generalized models for capturing the spread of such influential movements online based on #MeToo data. Such models may be applied to future hashtag movements and social and political discourse. In the digital age, with conversations happening more and more online, there is ample opportunity for independent actors to organize into a unified online movement and for people of different communities to interact, sparking conversation.

Epidemiological models of disease spreading in their simplest form study two populations, a susceptible and an infected population. This is known as the SI model. In this model, $S(t)$ is the number of susceptible individuals at time t and $I(t)$ is the number of infected individuals at time t. Each population is characterized by a differential equation $\frac{dS}{dt}$, and $\frac{dI}{dt}$ that represents how the number of individuals in that population changes over time [17]. In the context of idea spreading, the infected population represents people who are believers and advocates of a certain idea. The SI model is frequently extended into an SIR model by adding a recovered population of people $R(t)$ who were infected but are now free from infection. Other further extensions of the model include additional groups such as a passively immune population, an exposed population, or a population of people who adopt the infection differently (such as skeptics in the case of idea spreading) [7, 13, 16, 17]. Others have toyed with changing transmission and immunity parameters and examined their impact on infection [19, 27]. Works that have applied epidemiological models to the spread of ideas without taking into account the underlying structure of a social network have made the assumption of "homogeneous mixing," which means that each individual has an equal probability of transmitting or hearing an idea. In this case, the different state populations are only a function of time and some different transition probability parameters. This is a large simplifying assumption given the nature of idea spreading to be facilitated by social connections.

There has, however, been a variety of work that does not make this assumption and applies epidemiological models to the spread of ideas on a network structure. These works use the simplest epidemiological models and a degree block approximation to account for how nodes of varying degrees will act differently [4, 18, 22]. In addition to classical epidemic models, two other models including a threshold model and an independent cascades model have been used to define the spread of information on a network [11, 32]. Analysis done on network structure's impact on information diffusion and large scale coordination has found that scale free networks are poor for both while networks that have multiple tight communities and some inter-community connections are the most effective in facilitating the spread of information [23, 25, 33]. For the purpose of my model I choose a set network structure representative of the #MeToo movement and analyze the spread of "infection" for 3 different extended epidemiological models over that network.

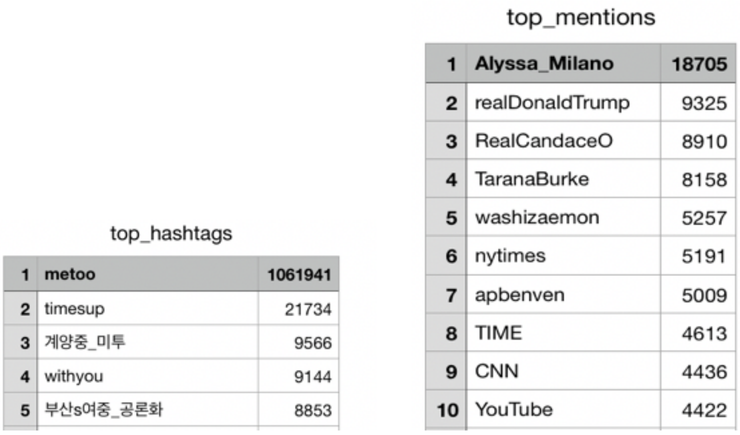
2 Background and Data

2.1 Overview

Activist Tarana Burke coined the phrase "Me Too" in 2006 to show solidarity to survivors of sexual assault. On October 15th, 2017 actress Alyssa Milano tweeted, "If you've been sexually harassed or assaulted write 'me too' as a reply to this tweet." By the following morning 55,000 people had replied, and the #MeToo movement as we know it today was born [26]. The data against which the following model will be tested is a sample of 10% of all of the tweets with #MeToo from October 16th, the day after the movement took off on Twitter, until December 14th, 2018. The sample features 1,434,996 tweets from 798,533 unique users during the time period.

The sample of tweets is in JSON format as a series of nested dictionaries where data is stored by keys. The following will help clarify some Twitter specific terminology. A *user* is a person or organization with a Twitter account that can send out tweets and is usually referred to by its screen name/user name. A user's *followers* are the people that follow that account and see the tweets posted by that user. A user's *friends* are the users that that user follows. A *hashtag* is a # symbol followed by a word that then tags the tweet as having that hashtag. A tweet can have as many hashtags as fit in the character limit.

There are many actions a user can do on Twitter. A user can tweet his or her own original tweet of 280 characters or less. A user can also share another user's tweet. This action is called a *retweet*. Retweets are commonly used as signs of affirmation. If a user agrees with the content of a tweet and wants to share it with his or her own followers, the user will retweet that tweet. Another sign of affirmation is a *favorite*. A user can favorite another user's tweet, showing that that user supports it and agrees with the content of the tweet. A different action a user can take is to *quote* a tweet. In doing this, a user shares the original tweet with his or her followers while adding his or her own comment on it. This is used either as a sign of affirmation or of contradiction. Similarly, a *reply*, where a user can comment directly in response to someone else's tweet, can be either a sign of affirmation or of contradiction. Much of the discourse that takes place between users with differing view points occurs as replies to tweets. If user A disagrees with the content of user B's tweet user A might convey this by replying to B's tweet, but if user A wants to make a public display of this disagreement, A will quote the tweet and express his or her disagreement with it for all of user A's own followers to see. Finally, a user may *mention* another user. This is when one user includes an @ sign followed by another user name in his or her tweet. Users will usually mention another user if they are calling on that user to do something or referring to something that user has done. The top 10 users mentioned and top 5 hashtags included in the tweets along with #MeToo are below.



top_hashtags

1	**metoo**	**1061941**
2	timesup	21734
3	계양중_미투	9566
4	withyou	9144
5	부산s여중_공론화	8853

top_mentions

1	**Alyssa_Milano**	**18705**
2	realDonaldTrump	9325
3	RealCandaceO	8910
4	TaranaBurke	8158
5	washizaemon	5257
6	nytimes	5191
7	apbenven	5009
8	TIME	4613
9	CNN	4436
10	YouTube	4422

Figure 1: The top 5 hashtags and top 10 user mentions appearing in the sample of #MeToo tweets.

2.2 Tweets Over Time

There are various days that are peaks in the number of tweets for a day over the course of the data set. The first is the first day for which I have data, Oct 16th, 2017, the day following Alyssa Milano's tweet. The other big spike occurs on September 9th driven by tweets about the movement in South Korea and the resignation of CBS's Leslie Moonves. The following period of high activity in late September and early October is driven largely by tweets about the controversy around then Supreme Court Justice nominee Brett Kavanaugh and the claim he sexually assaulted Christine Blasey Ford. Below is a graph of the log number of tweets each day over the duration of the data set. Spikes in the log number of tweets above one standard deviation about the mean are labeled with the corresponding events of the day.

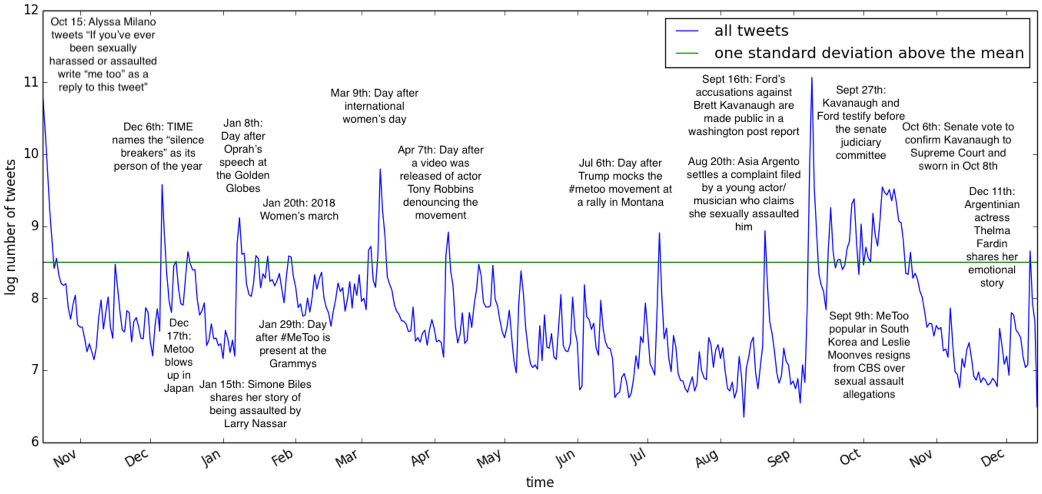


Figure 2: The log number of tweets per day over the duration of the data set. The green line is one standard deviation above the mean. Peaks above one standard deviation of the mean are labeled with the corresponding events.

A trend to note is that in the beginning of the data set, the events corresponding to peak days of aggregate numbers of tweets are generally positive events for the movement that are spreading information about the movement to people who may have not heard about it previously. Then starting with the April 7th spike in tweets, conversation on Twitter about #MeToo becomes much more controversial with famous individuals mocking and denouncing the movement and others speaking out to defend those who have been accused of sexual assault. This trend is consistent with Barbera's 2015 finding that controversial discourse on Twitter begins as a national conversation and then transforms into a "polarized exchange" [5]. The evolution from a conversation to a polarized exchange is somewhat captured in the data set by the slightly increasing fraction over time of tweets that are either quotes or replies. Despite a small increase in the tweet types that signal more conversation rather than complete agreement and support as does a retweet, retweets remain a significantly high fraction of all the tweets tweeted about #MeToo, averaging around 65%.

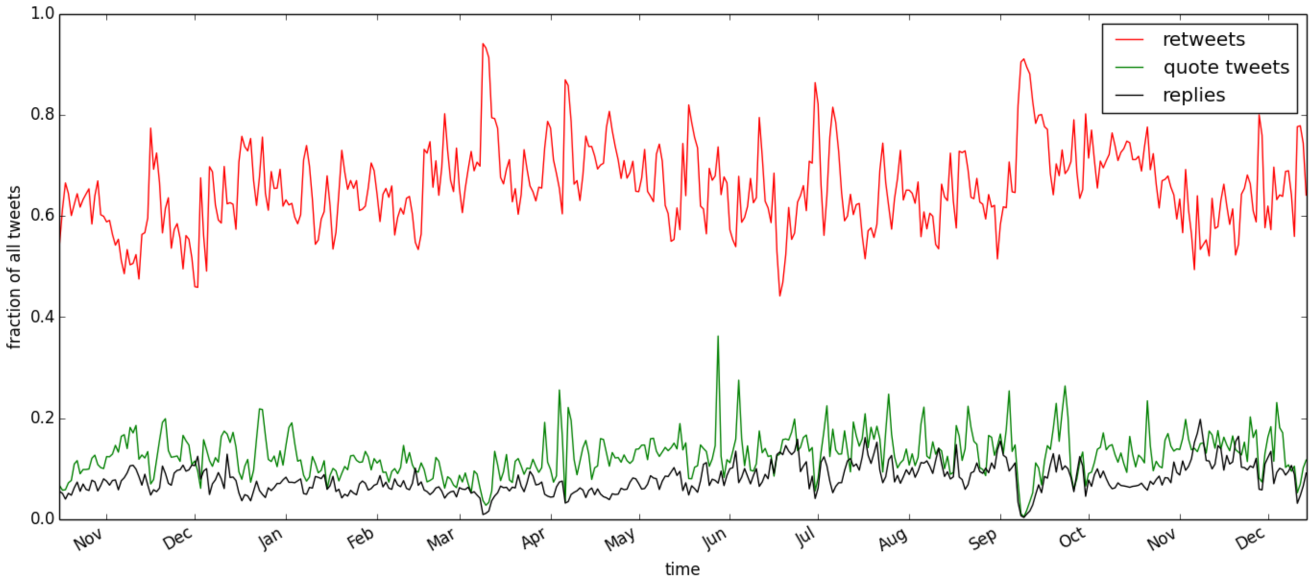


Figure 3: Fraction of tweets that are retweets, quotes, and replies

2.3 Identifying Groups

To begin to understand who is tweeting about #MeToo and how the conversation is spreading, I want to find communities that exist within the movement and the most influential members of those communities. Ideally I would be able to reconstruct a Twitter social network of followers and friends where users are represented by nodes and if one user follows another there is a directed edge drawn between them. This would allow me to see specifically for the #MeToo movement how likely a user is to tweet #MeToo given how many of that user's friends tweeted the hashtag. Unfortunately, the data set does not contain information on who each user's followers and friends are and thus this graph is impossible to construct. Instead, I create and explore a proxy for network structure and possible communities by looking at users' retweets. A retweet network is one in which each node represents a different user. A directed edge is drawn from node A to node B if the user represented by node A retweeted the user represented by node B. I created a visualization of this network for each of the 16 days in the data set with spikes in number of tweets above one standard deviation above the mean (see appendix).

To extract the community structure of the network for each popular day I use the modularity optimization algorithm presented by Blondel et al [8]. The algorithm has two parts that are repeated. The first assigns each node in the network its own community such that there are as many communities as nodes. Then, it looks at each node i and its j neighbors and evaluates the gain in modularity of taking node i and placing it into community j. Node i is then placed into the community which results in the maximum gain in modularity. The second part reconstructs the network such that its nodes are the communities that were formed during the first phase. The weights of the edges between new nodes are now the sum of the weights of the edges between nodes in the corresponding two communities. Once the second phase is completed, the first phase is again applied until a maximum modularity is obtained and no more changes can be made. Modularity is the fraction of the edges that fall within a community minus the expected fraction of edges between those same nodes if edges were distributed at random. Gain in modularity, ΔQ, is calculated as follows.

$$\Delta Q = [\frac{\Sigma_{in}+2k_{i,in}}{2m} - (\frac{\Sigma_{tot}+k_i}{2m})^2] - [\frac{\Sigma_{in}}{2m} - (\frac{\Sigma_{tot}}{2m})^2 - (\frac{k_i}{2m})^2]$$

Here Σ_{in} is the sum of weights of links within a given community. In the very first phase this is the sum of the number of edges since all edges have weight 1. Σ_{tot} is the sum of the weights of the links to the nodes neighboring the community. Similarly, k_i is the sum of the weights of the edges between node i and its neighbors and $K_{i,in}$ is the sum of the weights of the edges from node i to nodes in a given community. m is the sum of the weights of all the edges in the network.

The resulting communities are depicted by clusters, the larger of which are colored. The node with the highest in-degree, (i.e. the user that is retweeted the most on that day) is labeled in each community. Figure 4 below is an example of the retweet network from August 20th, 2018, the day that Asia Argento settled sexual assault accusations made against her by young actor Jimmy Bennett. The resulting popular users (who were retweeted the most) and tweets for that day reveal a wide diversity in users and in opinions regarding the news. There is large geographic diversity. @AnaPaulaVolei is a Brazilian pro volleyball player, @Brevesdepresse and @Melusine_2 are french accounts, @junktokyo is Japanese, @CristinaSegui is venezuelen, @el_pais and @alonso_dm are spanish, and @FionaBa47662575 is Australian. The remaining accounts are American. Within the American accounts, @CHSommers, @Danblackroyd, and @ajplus are liberal activists users while @LizCrokin is a conservative pundit. @WiredSources is a tech focused news source. And finally, @nytimes is a mainstream left leaning news outlet.

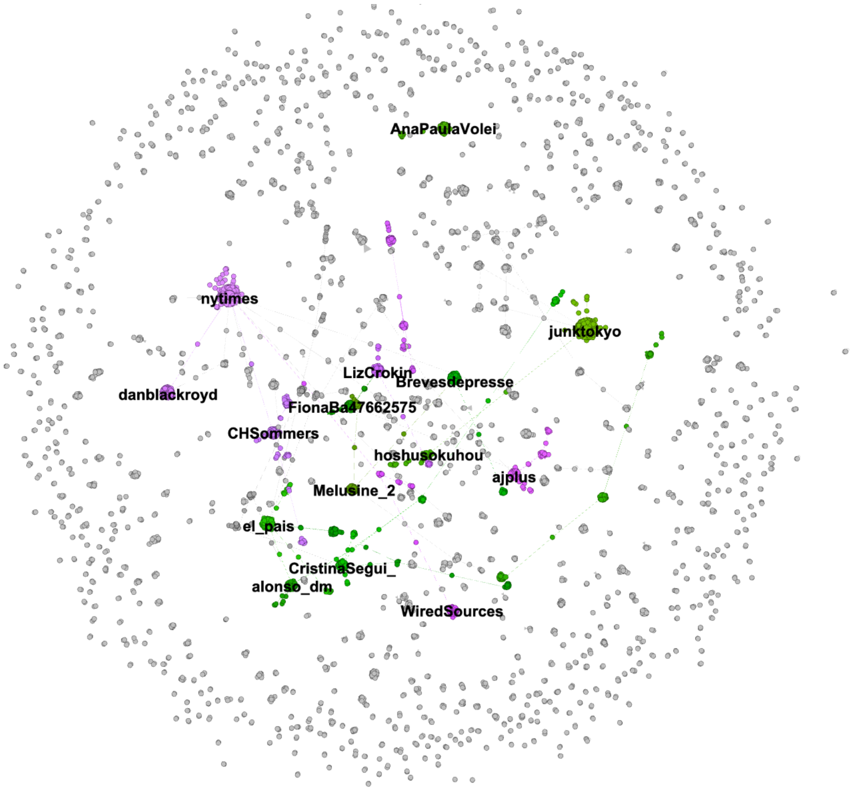


Figure 4: Sample retweet network from August 20th. Clusters in purple are people retweeting the labeled nodes that are representative of domestically located users. Nodes in green are users grouped around retweeting the labeled foreign users. Clusters around labeled nodes are hundreds of nodes tightly packed together.

Retweet networks for other days show a similar pattern of having users from various locations as well as users of varying opinions on the specific topic at hand and the movement as a whole. Unfortunately, given the nature of data it is difficult to tease out the later divide and place each and every tweet into a group. Ideally sentiment analysis could be done on tweets to figure out what the user's stance on the topic is and place that user into an ideological group. However, because tweets may have a positive sentiment while they support someone who is an activist very involved in #MeToo or someone who denounces the movement, sentiment of a tweet alone does not allow us to place users into ideological groups. It is possible though to place users into groups based on location. A fraction of the tweets in the data set contain information on the location of the tweet, and all tweets contain information on the language of the user. Below is a breakdown of the number tweets by their user's language.

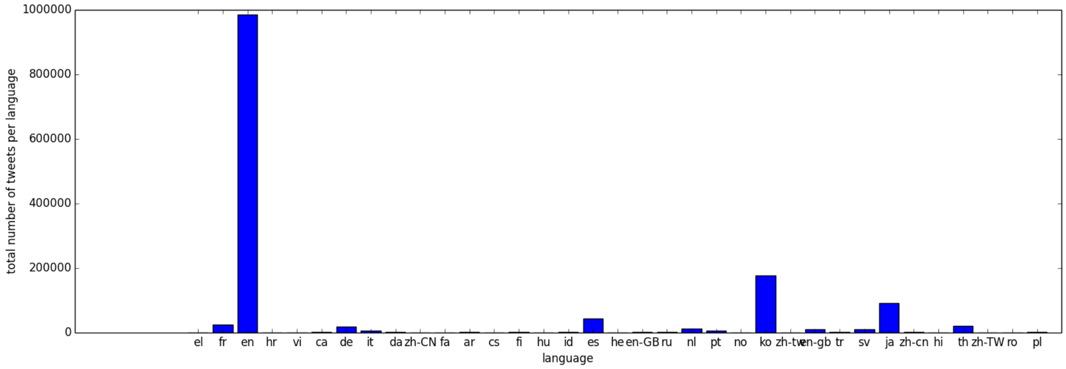


Figure 5: Break down of user languages by language code

Because every tweet has user data associate with it, I will use the user's language to classify tweets and corresponding users of those tweets into groups. Note that there is a distinction between English "en" and British English "en-gb" (see appendix for list of language codes) [10]. Thus, I make the assumption that most tweets in English are based in the United States and consider a domestic group of U.S. based users who tweet in English and an international group of users who tweet in some language other than English.

The international group includes British English. Below are graphs of the fraction of users and the total number of tweets broken down by group over the course of the data set.

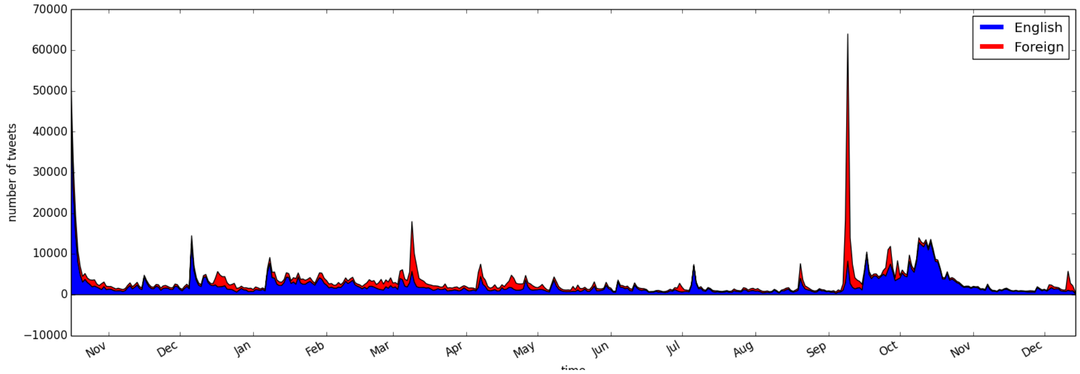


Figure 6: Number of tweets by group stacked

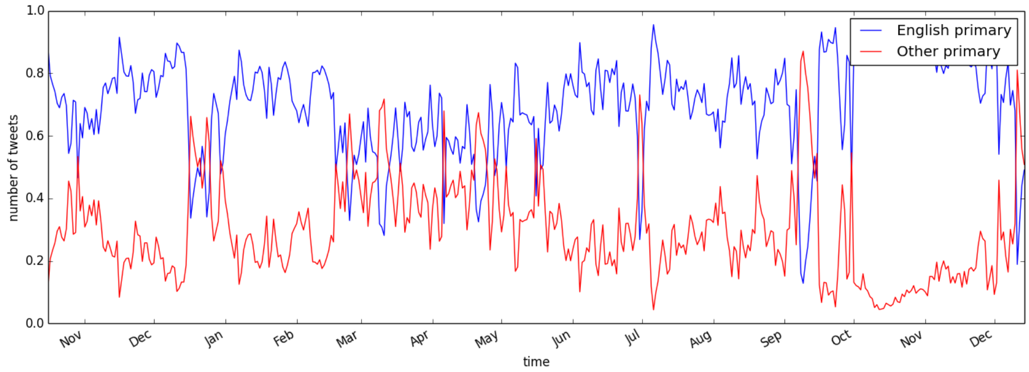


Figure 7: Percent of tweets by group

We see that the majority of tweets seem to be in the domestic group with a small contribution from the international group. As a whole there are 984,409 English tweets and 450,582 foreign language tweets. There are a few days when the international group tweets a lot. The four days during which international activity trumps domestic activity are December 17th 2017, when

#MeToo was most popular in Japan, March 9th, International Women's Day, September 9th, when the hashtag was very popular in South Korea, and December 11th 2018, when an Argentinian actress shared her story. Besides these days, conversation appears to be driven mostly by domestic issues with some international commentary on the events.

Because we care about movement's reach in terms of different individuals, we can think about the two groups not just by number of tweets but rather by number of distinct users involved. Below is a graph of the number of users in each group over the data set.

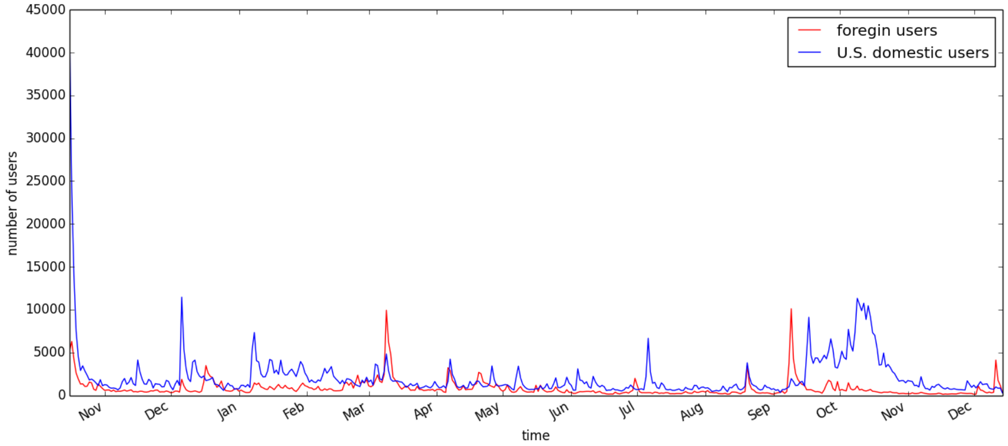


Figure 8: Number of users by day and by group

In aggregate of the 798,533 unique users in the data set, there are 566,626 English users (71%) and 231,907 foreign language users (29%).

2.4 User Entry and Group Dynamics

To better understand the spread of the popularity of the hashtag and the conversation around it, I explore the recruitment of new users into the data set and level of between group interaction. Below is a graph of new users per day over the data set and the fraction of all users that are new vs the fraction of all users that are repeat users over the data set.

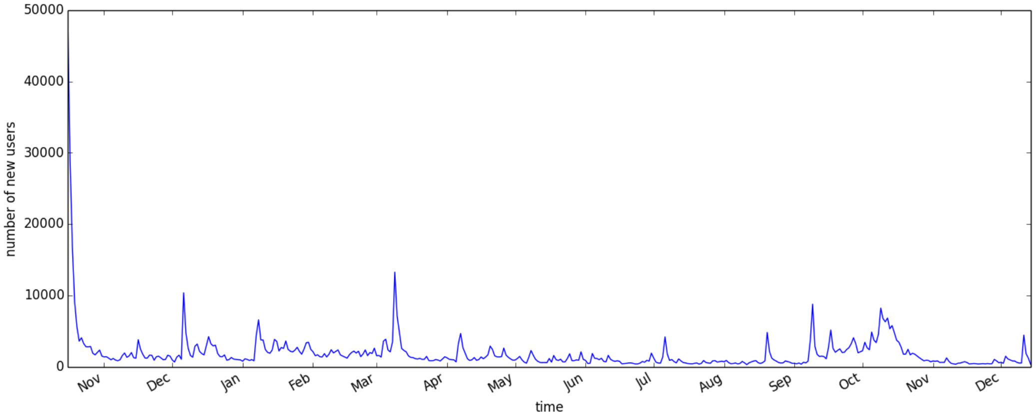


Figure 9: Number of new users per day over the course of the data set

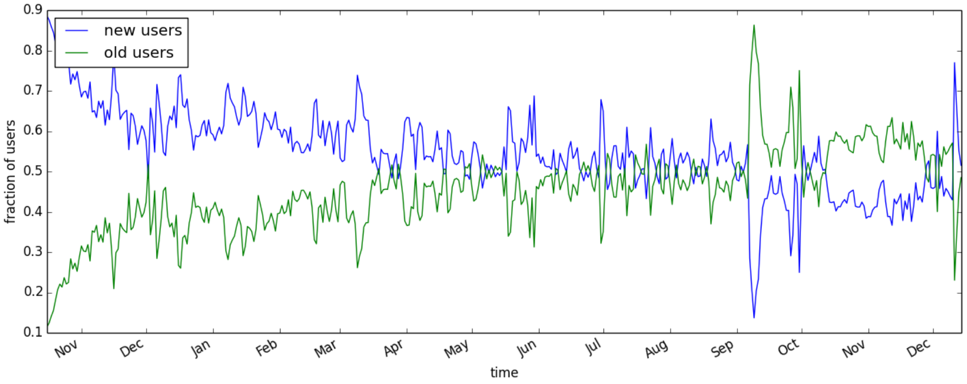


Figure 10: Fraction of new vs old users over the course of the data set

We see that the fraction of users who have appeared before in the data set steadily increases. In September 2018 the fraction of users who have appeared before in the data set increases to greater than 50% so that more users are repeat users than new users. This implies saturation in the spread over the social network so that the movement has reached most of the users that will tweet about it and is now being carried by users who are reactivating their use of the hashtag.

I now explore between group dynamics related to spread by considering how new users from each group are being drawn into the data set. As we saw earlier, retweets account for about 65% of all tweets at any time and they contain information on both the original tweeter and the retweeter. Thus, I use retweets to examine how users of different groups are being drawn into the data set. Below is a graph of the number of new users who enter the data set from retweeting someone not in their group.

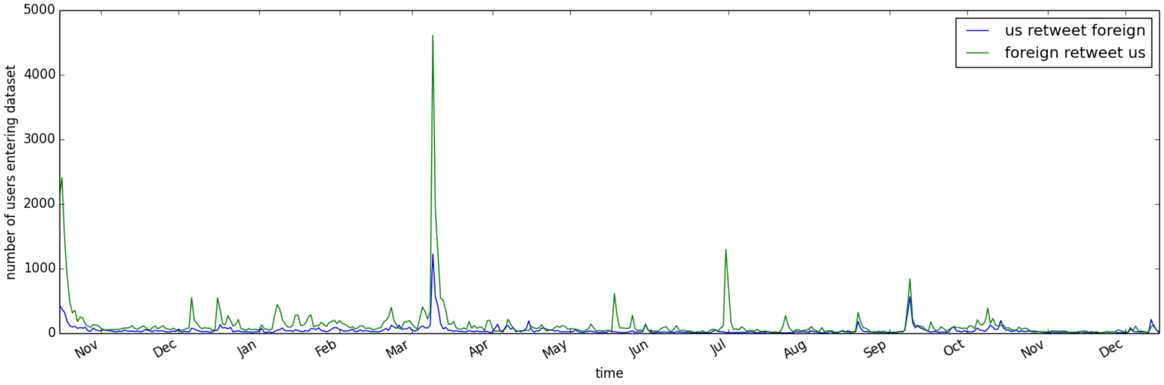


Figure 11: Number of new users entering from a cross-group retweet

In aggregate there are 350,634 U.S. users retweeting U.S. users and 19,441 U.S. users retweeting foreign users to enter the data set for the first time. There are 126,330 foreign users retweeting foreign users and 57,127 foreign users retweeting U.S. users upon their entrance into the data set. The ratio of international users retweeeting U.S. users to international users retweeting each other is about 0.45. The ratio of U.S. users retweeting international users, to U.S. users retweeting each other is 0.05. This indicates that the movement online surrounding #MeToo is largely U.S. driven. Overall, of the 553,532 users who enter the dataset on a retweet, 86% come from in group connections and 14% come from between group connections.

2.5 User Duration

Beyond thinking about large scale and group level dynamics, it is important to consider individual user activity and engagement with #MeToo to understand the dynamics of the spread of the movement. The movement and day to day popularity of the hashtag are driven by individual actors. To understand how active users are engaging with #MeToo I plotted raster plots of the days users were active. Below are two raster plots of individual user activity over the day count where day 0 is the first day in the data set and day 425 is the last.

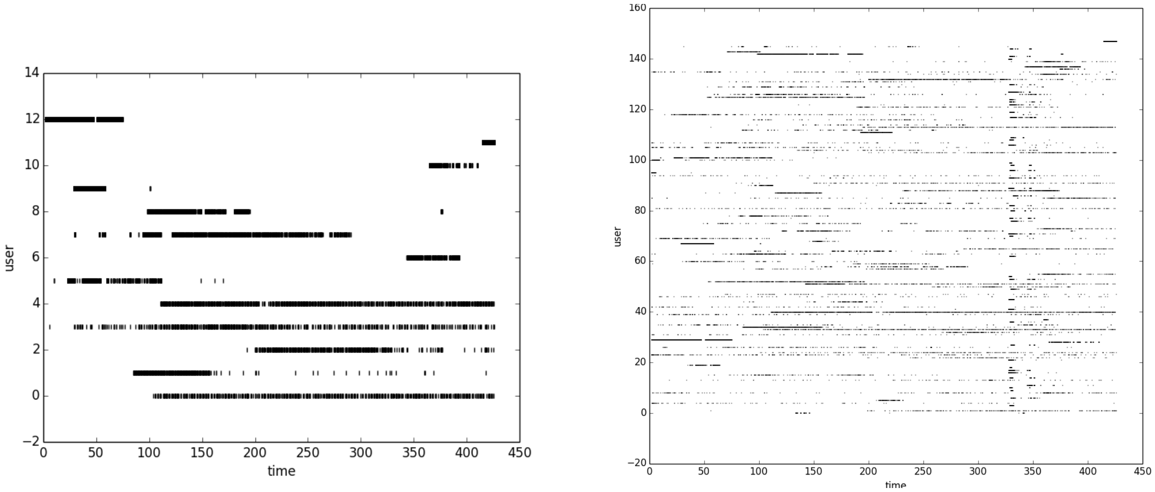


Figure 12: Raster plots of users who show up more than 500 times (left) and of users who show up more than 100 times (right)

It appears as though the most active users captured in these plots have periods of activity and inactivity such that they may be active for a short period of days at a time but have a longer span between their first and last days in the data set. A bootstrapped estimate from 80 sub-samples ranging from 1.3% to 98.7% of the data for the average number of days between bursts of activity is 131.15 with a standard deviation of 0.99. Let us call the short bursts of activity in which users show up on consecutive days a short duration. We then call the time between a user's first tweet and last tweet in the data set the user's long duration. Below are graphs of the number of users with the given short and long durations.

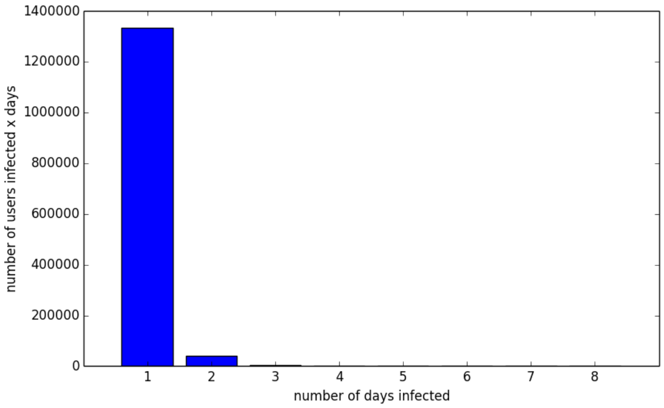


Figure 13: Short user duration where a duration of 1 means the user appeared in the data set for only one day during a given bust of activity.

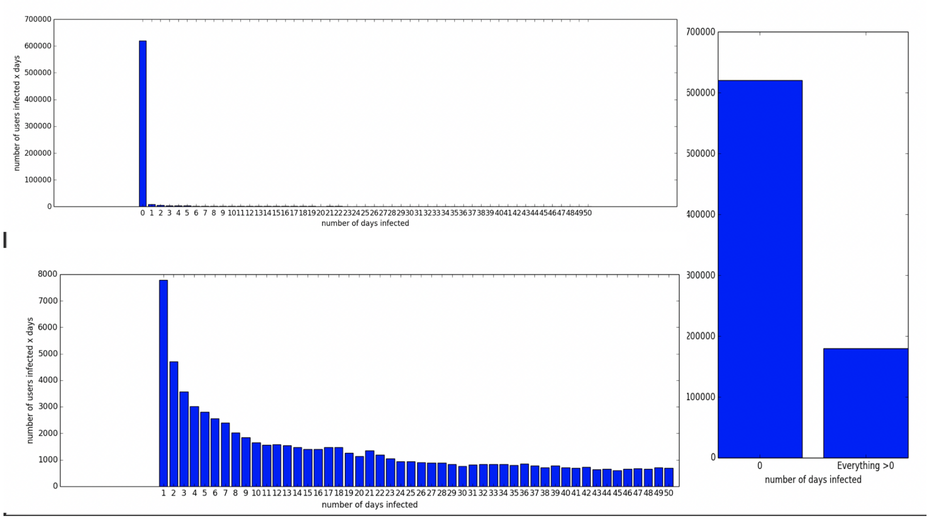


Figure 14: Long user duration in 3 forms. Here a duration of 0 implies being present in the data set on just one day and no others. Top left includes a duration of 0 and bottom left omits a duration of 0.

It is by far most common for users to show up only for one consecutive day at a time and to only show up once. 5 users show up for 8 consecutive days, the longest period of consecutive days that any user shows up in the data set. The average long duration between a user's first and last tweets follow a much larger distribution. Displayed on the graph are up to 50 day long durations but there are a total of 108,083 users with durations even longer than 50 days.

I bootstrapped the estimates for average short and long user durations using 80 sub-samples ranging from 1.3% to 98.7% of the data. The estimate for the average short duration of a given user is 1.0394 days with a standard deviation of 0.0008. The estimate for average long duration for a given user is 28.1060 days with a standard deviation of 0.1320.

2.6 Estimation of Total Number of Users Engaged

It is important to keep in mind that the above analysis of user duration and level of engagement using the data set is done on a 10% sample of tweets. We know then that we are potentially not capturing all of any given user's activity. People who show up only once in the data set could have only tweeted once in real life but likely tweeted some number of times more than that. Below is a graph of the number of users who appear a given number of times in the data set. There are users who appear more than 10 times, but the number who do becomes so small it can not be seen in comparison with the number who show up once.

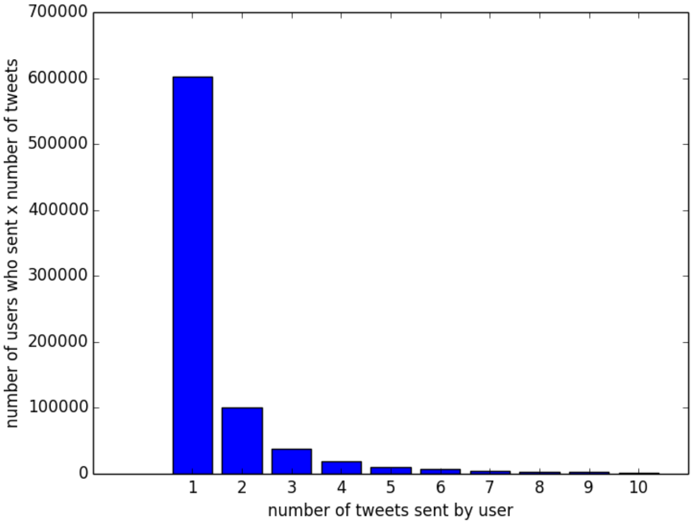


Figure 15: Number of users appearing to tweet a given number of times in our data set

A Twitter user that sends one tweet has a $\frac{1}{10}$ chance of showing up in the data set. A user that tweets twice has a $2 * \frac{1}{10} * \frac{9}{10} = \frac{18}{100}$ chance of showing up. This pattern continues such that the probability a user shows up once in the data set is binomially distributed with $P(X = 1) = \binom{n}{1} p^1 (1-p)^{n-1} = \binom{n}{1} \frac{1}{10}^1 (\frac{9}{10})^{n-1}$. Let P_n be the number of people who tweeted n times in real life and M be the max number of tweets from a single user. There are 602,362 users who appear once in the data set. Thus,

$$602,362 = \frac{1}{10} P_1 + \frac{18}{100} P_2 + \frac{243}{1000} P_3 + \cdots = \sum_{n=1}^{M} \binom{n}{1} \frac{1}{10}^{1} \frac{9}{10}^{n-1} P_n$$

There are 100,819 users who appear twice in the data set. So,

$$100,819 = \frac{1}{100} P_2 + \frac{27}{1000} P_3 + \cdots = \sum_{n=2}^{M} \binom{n}{2} \frac{1}{10}^{2} \frac{9}{10}^{n-2} P_n$$

This continues for all $k \in [1, M]$ where M is still the max number of tweets from a single user. So,

$$\text{Total Infected Population} = \sum_{j=1}^{M}\sum_{n=k}^{M}\binom{n}{k}\frac{1}{10}^{k}\frac{9}{10}^{n-k}P_n$$

An approximation of the true total population of Twitter users tweeting #MeToo can be found by solving for the P_n's and summing them together. I construct a matrix A of the coefficients and a vector from the data of the number of users appearing n number of times to solve for a vector of P_n's.

$$A = \begin{bmatrix} 1/10 & 18/100 & 243/1000\ldots \\ 0 & 1/100 & 27/1000\ldots \\ 0 & 0 & 1/1000\ldots \\ \vdots & \vdots & \vdots \end{bmatrix}, d = \begin{bmatrix} 602362 \\ 100819 \\ 37491 \\ \vdots \end{bmatrix}, p = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \end{bmatrix}$$

I want to solve for p in $Ap = d$ to find an estimate of the full population of Twitter users who tweet #MeToo at some point during the time of the data set. Unfortunately, the dimension of p is unknown. Because it is most likely that $dim(p) > dim(d)$, we have an under-determined system. I thus find the non-negative solution to the linear least squares problem, min $||A(p) - d||$. I take data I have for number of users appearing up to 55 times as my vector d and use it to solve for the first 660 P_n's. Summing these P_n's gives an estimate of 4,183,700 users that engage with #MeToo on Twitter.

To get a sense of sensitivity of this estimate I solve for p using 54 and 56 data values for the number of user appearances. Using the number of users appearing up to 54 times as the d vector yields an estimate of 4,183,600 users who engage with the hashtag on Twitter. And using the number of users who appear 56 times yields a 4,183,800 user estimate. This is a 0.002390228% change in both directions. I compute the population estimate using anywhere between 10 data values and 84 data values, the maximum number of data values for which numbers of user appearances are consecutive. The range of estimated populations is narrow, from about 4.1 million to about 4.3 million users.

2.7 Exogenous Factors

When thinking about social movements associated with hashtags I make the distinction between the popularity of the hashtag itself at a given time and the strength of the overall movement. A hashtag may be used more on a specific day because of a unique event. That event may even be exogenously induced. For example, there was a spike in the number of tweets with #MeToo during the period of time that Brett Kavanaugh was on trial. This event was covered largely by mass media outlets beyond Twitter rather than by the individuals involved tweeting themselves. Thus, a lot of users during this period may have been inspired to tweet as a result of exogenous sources rather than someone immediately in their social network. Though there is a distinction between exogenous and endogenous factors when thinking about the popularity of the hashtag, all of these factors (the news, the tweets, the overall sentiment) are endogenous to the movement itself. In an attempt to quantify the exogenous effect of the mass news outlets in the social network, I examine the number of tweets per day coming from news accounts. Though the news accounts themselves are in the social network, I use this as a proxy to understand the influence the news has on information spreading. I defined Twitter news accounts by compiling a list of the most popular news outlets' Twitter accounts [29, 2]. Below is a graph of the fraction of tweets from and retweets of news accounts on Twitter compared with the log number of total tweets.

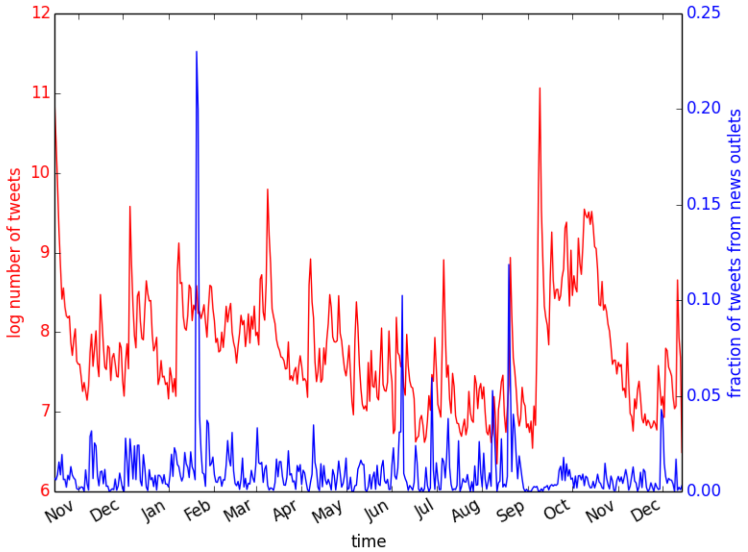


Figure 16: Log number of tweets in red and fraction of tweets originating from news sources in blue

The average percent of all tweets that are either tweets from a mass news source's Twitter account or retweets of that account is 1.09%. The fraction of tweets coming from news accounts appears to vary significantly, anywhere between 0 and 20%. The day with the single largest spike in mass media coverage in Jan 20th, 2018, the day of the 2018 Women's March.

The estimate that about 1% of users are tweeting in response to exogenous news sources is subject to two serious limitations. First, I chose to track only the most followed Twitter news accounts and there are more news accounts that exist. Second, this data does not capture users influenced to tweet by news mediums other than Twitter. As a result, the estimate for the number of tweets from news sources is likely too small.

3 Model Construction

3.1 A Hashtag vs. A Movement

Though Tarana Burke created the idea of MeToo back in 2006, the movement as we know it today began as a reaction to Alyssa Milano's tweet in 2017. The tweet sparked viral popularity of #MeToo. The hashtag has continued to spike in popularity since the movement began, again reaching viral levels as it spreads to new countries and major events occur. However, after the first viral spike in popularity, the movement has expanded beyond the hashtag. Over 200 men have lost their jobs or rolls as a result of #MeToo accusations and nearly half have been replaced by women [3]. Not all of this was done online through tweets accompanied with the hashtag. Thus, it is important to distinguish between the popularity of the hashtag and of the movement as a whole when modeling its dynamics.

In the following section I use three different versions of a model each with susceptible, infected, and recovered classes of people where infected can be interpreted to mean engaging with MeToo. When modeling the popularity of the hashtag I consider a user to be infected on a day if he or she actively tweets #MeToo at least once on that day. I call this type of infection “actively infected.” When modeling the movement as a whole I consider an individual to be infected if he or she is still broadly thinking about and engaging with the movement over a time period. An individual might only tweet or retweet #MeToo every once in awhile, but in the interim still reads news articles, discusses with friends, and espouses beliefs about the movement that help it to spread and gain momentum. I try to approximate this period of broader engagement with the movement by considering a user to be infected from his or her first tweet with #MeToo in it to the last. I call a user with this type of infection “still infected.”

3.2 Basic Models Introduction

I examine three different models and their effectiveness at capturing the popularity of #MeToo and the movement as a whole. They are a basic SIR model, and two extensions, SISR and SIRS. First let us consider these dynamical systems under the assumption of homogeneous mixing. This means that any one individual can transition between states at any time with some probability.

The following models rely on the parameters β, γ, ξ and assume a fixed population N, where $N = S + I + R$, the populations of susceptible, infected, and recovered individuals respectively. Though users leave Twitter and new users enter every day, I assume that the population of Twitter users is large enough that this will not significantly impact the dynamics of the spread of the hashtag or the movement. So, I choose not to account for these vital dynamics.

A basic SIR model is one in which each individual, if it infected, will eventually become recovered and the infection will die out after some period of time. The infection may die out before infecting all individuals in the susceptible population. The period of time of infection is determined by β, and γ, infection and recovery probabilities respectively.

Figure 17: Schematic of an SIR model, with transition parameters β and γ

A SIR model captures how populations $S(t)$, $I(t)$, $R(t)$ change over time. This change is described by the following set of differential equations where $N = S + I + R$ and $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$.

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

An SISR model is one in which infected users either become recovered or directly become susceptible again. If a user becomes susceptible again he or she can become infected again, but if a user becomes recovered he or she will remain recovered. Here ξ is the transition probability of going from the infected group back to the susceptible group.

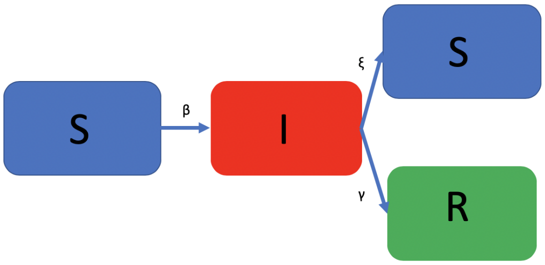


Figure 18: Schematic of an SISR model with transition parameters β, γ, and ξ

A SISR model still captures how populations $S(t)$, $I(t)$, $R(t)$ change over time where $N = S + I + R$ and $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$, but with slightly different dynamics. It is determined by the following differential equations. People in the infected population either leave to the recovered population with probability γ or back to the susceptible population with probability ξ.

$$\frac{dS}{dt} = -\beta SI + \xi I$$
$$\frac{dI}{dt} = \beta SI - \gamma I - \xi I$$
$$\frac{dR}{dt} = \gamma I$$

The final model is a SIRS model. In an SIRS model, infected users become recovered after some time. Then users that are recovered will once again become susceptible after some time. This generally creates cyclical behavior where as in the other two models, the infection eventually dies out for good. Note here that ξ is defined different than it is in the SISR model. Here ξ is the transition probability of going from the recovered group back to the susceptible group. This makes the time delay until a user gets reinfected different than the same delay in the SISR model because in SIRS the length of the time until reinfection depends on β and ξ rather than just β.

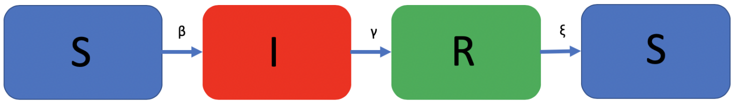


Figure 19: Schematic of an SIRS model with transition parameters β, γ, and ξ

A SIRS model is described by the following differential equations again where $N = S + I + R$ and $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$.

$$\frac{dS}{dt} = -\beta SI + \xi R$$
$$\frac{dI}{dt} = \beta SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I - \xi R$$

3.3 Model Populations in the Data

3.3.1 Susceptible

The data set does not give specific information on the susceptible population because the only users who show up are users I consider to be infected. Conceptually, the susceptible pool in these models is really the number of daily active Twitter users because I am examining the spread of infection on a day to day basis. Twitter recently published their number of DAUs for the first time in its Q4 2018 earnings report (previously they reported monthly active users). As of Q4 2018 there were 126 million DAUs on Twitter, 27 million of which are in the United States and 99 million of which are international [31]. This means about 20% of DAUs are domestic to the U.S. and about 80% are international. The report also has the number of DAUs over the past two years. The period during which #MeToo was popular on twitter saw a range of 114 million DAUs to 126 million DAUs [31].

Assuming a full infected population of about 4.2 million DAUs (from the previous calculation of the total number of people tweeting #MeToo) and 126 million DAUs, we see that about 3% of the entire twitter network gets infected with #MeToo at some point.

3.3.2 Infected

Because there is a distinction between modeling the movement and the popularity of the hashtag, there is a corresponding distinction in the way infection is quantified. Below are graphs of the infected population of users in the data set with infection defined in two different ways. First, infection in terms of actively infected users tweeting day by day gives us an idea of the popularity of the hashtag over time.

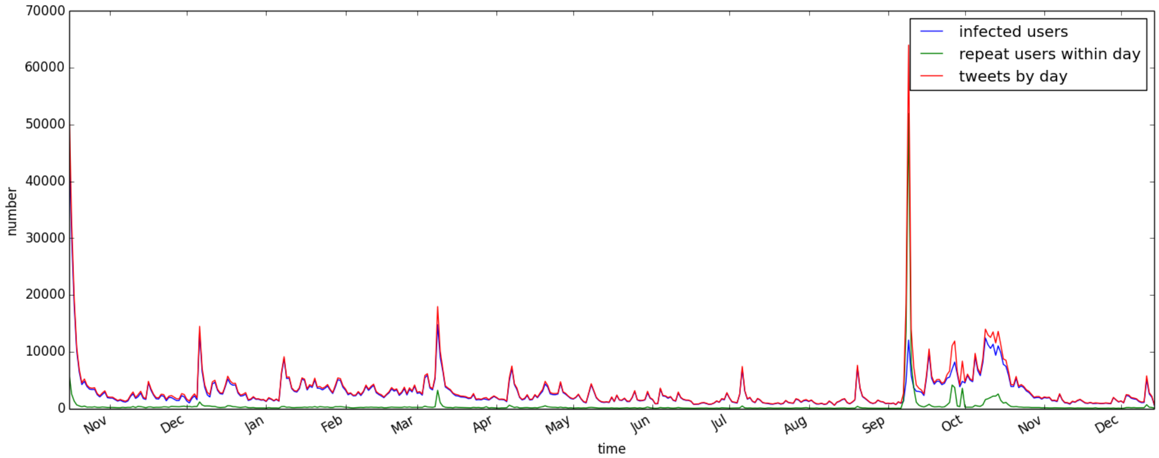


Figure 20: Actively infected users, representative of the popularity of the actual hashtag

The number of times a hashtag is tweeted is a result of the number of users tweeting the hashtag and the number of times a user tweets. The time scale of days which I choose to work with in the following models allows us only to consider the number of users who are infected on a day and gives no information about the number of tweets each user sends on that day. I include the total number of tweets as well as the number of repeat users on a day in the graph above to show that the number of active users generally very closely matches the number of tweets. This means the extended epidemiological models not only capture popularity of the hashtag among users but also reasonably represent the number of times the hashtag shows up on a day. The only day when this is not true is September 9th, when #MeToo blows up in South Korea. The large number of tweets on this day is driven mostly by users who show up multiple times on this single day.

The various explosions of popularity of #MeToo of this nature in foreign countries are to be expected given 80% of the Twitter network is international. However, of the infected users in our data set only about 30% are international. It is important to note that hashtags like #yotambien and #balancetonporc that show up in the data set are translations of #MeToo and likely show up on their own without the #MeToo tag that pulls them into our data set. This makes it possible that more than 30% of the whole movement was internationally driven though it is clear the movement began in and is dominated by U.S. users. The U.S. has a higher density of infection as about 70% of the infected population is coming from just about 20% of the entire Twitter network.

The second version of infection is users who are still infected as defined by a period between a user's first and last tweets. This version of infection gives us an idea of the popularity of the MeToo Movement over time. Below is a graph of the number of users who are still infected.

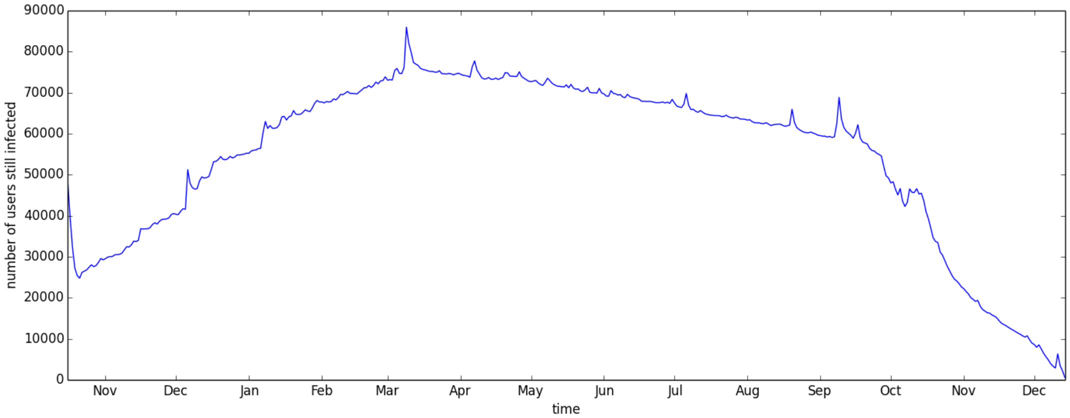


Figure 21: Still infected users, representative of user interest in the movement as a whole

Note that it is very possible users are still engaged with the movement as a whole but just do not show up in the data set again. This is likely as the MeToo Movement is still relevant today but the data set ends in December 2018, and thus is a limitation of the data set. Over a very long time horizon, however, it can be expected that the Me Too Movement at least by its name will loose momentum. When this happens the still infected population will exhibit a decline in infection similar to the period September to December of 2018 in the later graph but perhaps more gradual.

Another limitation of the data that is important to address is that we have only a 10% sample of all tweets with #MeToo. This means that users in actuality may tweet many more times than they appear. Thus, the total number of infected users will be much larger, and it is possible that the close to 1-to-1 relation between users and number of tweets observed in figure 20 is actually skewed.

3.3.3 Recovered

Because infection is defined in two different ways, the recovered population of users should be different for each definition of recovered. Below are two graphs. First, of users' short recovery after being actively infected and second of users' long recovery after the period in which they were still infected. Users recovered after being actively infected may once again become actively infected and enter the data set, but users recovered after their long, still infected period do not reenter the data set.

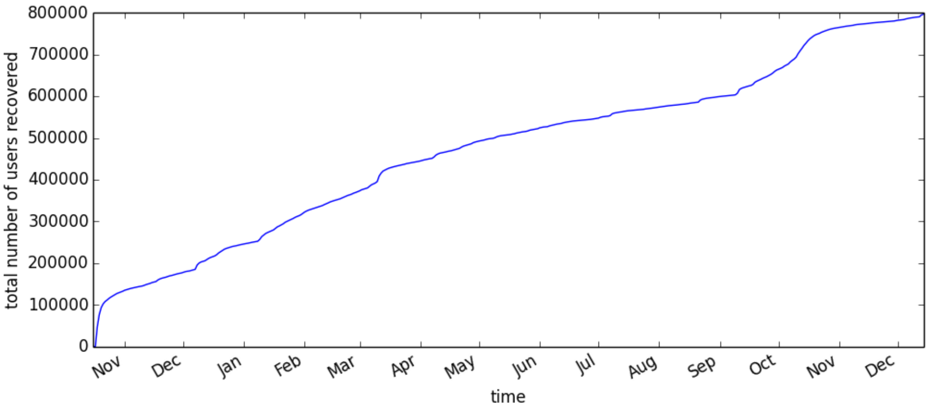


Figure 22: Short recovered curve as defined by short user duration being over allowing users to become infected once again

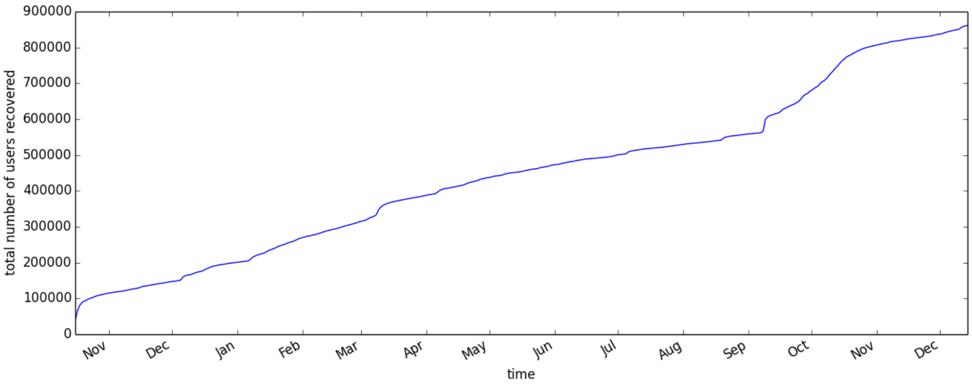


Figure 23: Recovered users as defined by a user's long duration being over, representative of people recovering from engaging with the movement as a whole

The shape of the short recovery curve looks a lot like the shape of the long recovery curve for two reasons. First, the vast majority of users are only infected once and thus the shape of the curve is largely determined by them. Second, even if users get reinfected it is most likely for only a day and so a few users exiting the recovered population for only a day at a time does not drastically change the shape of the curve.

3.4 Transition Parameter Estimation

3.4.1 Gamma

To estimate γ, the probability of transition from I to R in SIR and SIRS we must consider user durations. We can solve for γ from the average user duration. The probability a user transitions to recovery after the first day of infection is γ and the probability that user stays infected is $1 - \gamma$. The probability of a user's transition out of the infected population and into the recovered population on a certain day is as follows.

Day 1: γ

Day 2: $\gamma(1 - \gamma)$

Day 3: $\gamma(1 - \gamma)^2$

Day 4: $\gamma(1 - \gamma)^3$ and so on.

Then we can find the expected value of days infected before transitioning to recovered based on the parameter gamma. We do this taking the first moment.

Let $(1-\gamma) = p$

$$E = \sum_{n=1}^{\infty} n\gamma p^{n-1} = \gamma \sum_{n=1}^{\infty} np^{n-1}$$
$$\text{Let } S = \sum_{n=1}^{\infty} np^{n-1} = 1 + \sum_{n=2}^{\infty} np^{n-1}$$
$$\text{Then, } pS = \sum_{n=1}^{\infty} np^{n} = \sum_{n=2}^{\infty} (n-1)p^{n-1}$$
$$\text{So, } S - pS = 1 + \sum_{n=2}^{\infty} np^{n-1} - \sum_{n=2}^{\infty} (n-1)p^{n-1} = 1 + \sum_{n=2}^{\infty} p^{n-1}$$

Now to find $\sum_{n=2}^{\infty} p^{n-1}$ let, $R = \sum_{n=2}^{\infty} p^{n-1}$

$$\text{Then, } pR = \sum_{n=2}^{\infty} p^{n} = \sum_{n=3}^{\infty} p^{n-1} \text{and } R - pR = \sum_{n=2}^{\infty} p^{n-1} - \sum_{n=3}^{\infty} p^{n-1} = p$$
$$\text{Thus, } R = \frac{p}{1-p}$$
$$\text{Then, } S - pS = 1 + \frac{p}{1-p} \text{ and } S = \frac{p}{(1-p)^2} + \frac{1}{(1-p)}$$
$$E = \gamma S = \frac{1-\gamma}{\gamma^2} + \gamma\frac{n}{\gamma} = \frac{1-\gamma}{\gamma} + 1 = \frac{1}{\gamma}$$

Thus, the expected number of days a user is infected before transitioning, the average user duration, is $\frac{1}{\gamma}$. The average long user duration for still infected individuals in the data is 28.1060 days and the average short user duration of consecutive actively infected days is 1.0394. So for actively infected users $\gamma = \frac{1}{1.0394} = 0.96$ and for still infected users $\gamma = \frac{1}{28.1060} = 0.036$.

We can calculate 95% confidence intervals of these estimates.

$$95\% \text{ CI of Short } \gamma = \frac{1}{1.0394 \pm 1.96 * 0.0008} = [0.0960, 0.0964]$$
$$95\% \text{ CI of Long } \gamma = \frac{1}{28.1060 \pm 1.96 * 0.1320} = [0.035, 0.036]$$

These estimates for γ are reasonable for both actively infected and still infected users in the SIR model (i.e. in both the hashtag and movement contexts). The estimate of γ for actively infected users also makes sense in the SIRS model. However, due to the cyclical nature of the SIRS model, users always reactivate and thus we do not logically have a γ that means users become permanently recovered from the movement. For SISR, we must think about gamma differently.

In SISR on any given iteration users in the infected population can transition to the recovered population with probability γ. If they do not, they can become susceptible again with probability ξ. If a user does neither of these two then he or she will remain in the infected population for another day. Thus, for each iteration $\gamma + \xi \leq 1$. A user has the following probabilities of transitioning to the recovered class by iteration. One iteration is representative of one day.

Iteration 1: γ

Iteration 2: $\gamma(1-\gamma)(1-\xi)$

Iteration 3: $\gamma((1-\gamma)(1-\xi))^2$ and so on.

Thus, the total probability of transitioning to R is $\sum_{n=0}^{\infty} \gamma((1-\gamma)(1-\xi))^n$.

Let $x = (1-\gamma)(1-\xi)$. Then,

$$\sum_{n=0}^{\infty} \gamma((1-\gamma)(1-\xi))^n = \gamma \sum_{n=0}^{\infty} x^n = \gamma \frac{1}{1-x} = \frac{\gamma}{1-((1-\gamma)(1-\xi))} = \frac{\gamma}{\gamma+\xi+\gamma\xi}$$

Using similar logic, the total probability of transitioning to S is $\sum_{n=0}^{\infty} \xi(1-\gamma)((1-\gamma)(1-\xi))^n$. Let $x = (1-\gamma)(1-\xi)$. Then,

$$\sum_{n=0}^{\infty} \xi(1-\gamma)((1-\gamma)(1-\xi))^n = \xi(1-\gamma) \sum_{n=0}^{\infty} x^n = \frac{\xi(1+\gamma)}{\gamma+\xi+\gamma\xi}$$

We see that $1 - \frac{\gamma}{\gamma+\xi+\gamma\xi} = \frac{\xi(1+\gamma)}{\gamma+\xi+\gamma\xi}$. Thus, the complement of the probability of transitioning to R is the probability of transitioning to S. This makes sense as summing over all iterations, every user must transition from I to somewhere.

The total probability of transitioning to R can be approximated from the data by looking at the ratio of the number of users in the data set for one consecutive period to the total number of users. The number of users who appear for only one consecutive period in the data set is 608,199, and 190,334 users show up for more than one period. There are 798,533 total unique users.

$$\text{Total probability of transition to R } \approx \frac{608199}{798553} \approx 0.76 \approx \frac{\gamma}{\gamma+\xi+\gamma\xi}$$

Unfortunately, it is not possible to solve for γ and ξ independently with the information we can pull from the data set, but we know for an SISR model they are approximately related in the above way.

3.4.2 Beta

To estimate β, the probability of transitioning from S to I, we must think about β for the hashtag different than β for the movement. Both the beta from the data for the popularity of the hashtag and for the movement can be approximated by looking at the timescale until infection. As seen in the calculation of gamma above, when a user only has one possible population to transition to, the expected time in the prior population before transition is 1/(the relevant transition parameter). Because in all models users transition from a susceptible population only into an infected population, the time until infection is $\frac{1}{\beta}$.

To find β for the hashtag we can consider the time until the first peak infection. This infection is the largest in the data set. Because Alyssa Milano tweeted her famous tweet the day before the data set begins, the first big spike on the first day of the data set is really one day after the infection began. A limitation of using days as a time scale for the model means that time until peak infection implies $\beta = 1$. This would mean everyone would automatically become infected and cannot be true. Instead, I look at the total number of days in the first big infection spike. Including the day Alyssa Milano tweeted, the first big spike in infection lasts about 7 days before the next little spike in popularity. This implies $\beta = \frac{1}{7} \approx 0.14$. This estimate for β can logically be applied to SIR, SISR, and SIRS models of the popularity of the #MeToo.

The β for the movement as a whole can be derived by looking at the average time until first infection for any given user. I bootstapped an estimate for the average time until first infection with 80 sub-samples ranging from 1.3% to 98.7% of the data. The resulting estimated time period is 173.085 days with a standard deviation of 0.206. This estimate is equal to $\frac{1}{\beta}$.

$$95\% \text{ CI of } \beta = \frac{1}{173.085 \pm 1.96 * 0.206} \approx [0.00576, 0.00580]$$

This estimate of β is reasonably applied to the SIR and SISR models for the movement as a whole. This estimate of β does not make a lot of sense in an SIRS model applied to the movement a whole and in fact creates a contradiction in the estimation of ξ as explained in the following section.

3.4.3 Xi

We know ξ in the SISR model is related to γ, but let us consider ξ in the SIRS model. As seen in the calculation of gamma above, when a user only has one possible population to transition to, the expected time in the prior population before transition is 1/(the relevant transition parameter). Unfortunately, it is not possible to get information about the transition from R to S, the transition ξ controls, in the data. We do know, however, that the estimate of the average period between sets of consecutive active days of infection for a given user is 131.15 with a standard deviation of 0.99.

$$95\% \text{ CI of } \frac{1}{\xi} + \frac{1}{\beta} = 131.15 \pm 1.96 * 0.99 = [129.21, 133.09]$$

To go from one recovery to another infection, a user must transition from R to S, and S to I. We can thus expect the average time between infections to be equal to $\frac{1}{\xi} + \frac{1}{\beta}$. We have found estimates for β and can solve. First, I use the β estimate for the hashtag to find the ξ in SIRS that models the popularity of #MeToo.

$$\frac{1}{\beta} + \frac{1}{\xi} = 131.15$$

$$\frac{1}{\beta} + \frac{1}{\xi} - \frac{1}{\beta} = 131.15 - 7 = 124.15$$

So, $\xi \approx 0.008$

and $\xi \in [\frac{1}{133.09 - 7}, \frac{1}{129.21 - 7}] = [\frac{1}{126.09}, \frac{1}{122.21}] = [0.0079, 0.0082]$

Now we try the same process to find ξ in SIRS for the movement. In this case we know $\frac{1}{\beta} = 173.085$. Thus we see the following.

$$\frac{1}{\beta} + \frac{1}{\xi} = 131.15$$

$$\frac{1}{\beta} + \frac{1}{\xi} - \frac{1}{\beta} = 131.15 - 173.085 = -41.935$$

We see here that ξ would be a negative number. Even using estimates for $\frac{1}{\beta}+\frac{1}{\xi}$ in the upper range of its 95% confidence interval will not return a positive ξ. Thus, it is clear that the β estimated for the data for the movement does not work in an SIRS context. Because SIRS is cyclical, it struggles to capture the single long infection that occurs as users are engaging with the Me Too movement as a whole.

3.4.4 Summary

Despite the limitation that the data set only contains infected users, I have estimated data informed parameters for the three dynamical systems models. They are summarized in the table below. The bootstrapped parameter estimates have very small error ranges indicating that these parameter estimates are pretty sound in the data. These are the parameter estimates for the three models assuming homogeneous mixing. In its basic form with this assumption, these parameters control the entire model of infection. Beyond parameters, two other important factors that effect how the model captures infection are network structure and the choice of model framework.

Model	Hashtag	Movement
SIR	$\beta \approx 0.14$, $\gamma \in [0.0960, 0.0964]$	$\beta \in [0.00576, 0.00580]$, $\gamma \in [0.035, 0.036]$
SISR	$\beta \approx 0.14$, $0.76 \approx \frac{\gamma}{\gamma+\xi+\gamma\xi}$	$\beta \in [0.00576, 0.00580]$, $0.76 \approx \frac{\gamma}{\gamma+\xi+\gamma\xi}$
SIRS	$\beta \approx 0.14$, $\gamma \in [0.0960, 0.0964]$, $\xi \in [0.0079, 0.0082]$	$\beta =$N/A, $\gamma =$N/A, $\xi =$ N/A

3.5 Model Simulations on a Network

I consider SIR, SISR, and SIRS models not just as dynamical systems alone, but rather as dynamical systems run over a network. Thus I do not make the typical assumption of homogeneous mixing made by the basic epidemiological models. The following analysis comes from simulations I have constructed in python using an underlying network that is representative of Twitter's social network. The representative network is such that nodes symbolize users, and edges drawn between nodes represent users that follow each other on Twitter. I used the python package Network X's random partitioned graph generator to generate the graph underlying the simulation because it allows me to specify the number of groups, number of nodes in each group, and probabilities of a node sharing an edge with a node in its group and with a node in the opposite group. I chose a network with 1000 nodes and two main groups to represent the international and domestic breakdown. For this ratio I use the fact that Twitter DAUs are about 20% domestic and 80% international [31]. I choose a network with 1000 nodes because of computing restraints when running my simulation.

I arbitrarily choose the probability of nodes connecting within group to be 10%. I chose the probability of nodes connecting between groups to be one eight of that as in the #MeToo data, the number of users entering on a retweet of someone in the opposite group was about an eight of the number of users entering on a retweet of someone in their group. Below is a visualization of this underlying network on which I ran simulations. The green group represents the U.S. domestic Twitter users and the pink group is internationally located Twitter users.

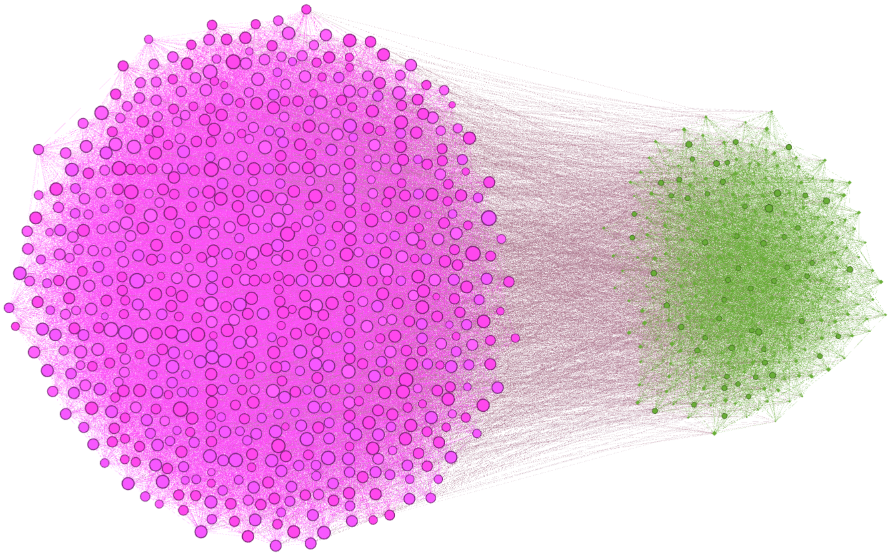


Figure 24: Visualization of network over which simulations are run

The basic premise of each model simulation I run is that one node in the network is randomly chosen to start as infected and all others are susceptible. I iterate through each neighboring node connected to the infected node and determine with a probability, β, if that neighbor will become infected. With each iteration more nodes will be infected and the iterating process will continue. In all simulations I run 426 iterations, representative of the 426 days in the data set. A node infected for an iteration represents a user that is actively infected (i.e. tweets on that day) for a day in the data set. Once a node is infected, there is a probability, γ, that the node will become recovered on any iteration. The two extensions beyond a basic SIR model also contain a probability parameter, ξ, that a node will again become susceptible and thus can be reinfected.

4 Model Evaluation

4.1 Standards For Evaluation

In this section I seek to evaluate two things, the optimality of parameter estimates and the ability of the model as a whole to capture the dynamics of MeToo. I examine each model type and its parameter estimates quantitatively and qualitatively with respect to the popularity of #MeToo and of the movement as a whole. I look at the populations over time, implied user durations, and the mean squared error of the model with a variety of metrics in the data.

I choose to examine the mean squared error between 3 pieces of information drawn from the data and the analogous pieces of information from the models. First, I minimize the mean squared error with respect to the transition probability parameters for each model. This allows us to evaluate the optimality of the parameter estimates in capturing key metrics from the data with the model. It also allows us to perhaps find more optimal parameters. In theory the model with the lowest mean squared error at the optimal parameters should be the best model to capture either the movement or the hashtag. I minimize mean squared error with respect to 3 metrics as follows.

$$
\begin{aligned}
\min_{\beta,\gamma,\xi}\ [(D1 - m1(\beta,\gamma,\xi))^2 + &(D2 - m2(\beta,\gamma,\xi)^2) + (D3 - m3(\beta,\gamma,\xi)^2)] \\
&= \min_{\beta,\gamma,\xi} E[(D - m(\beta,\gamma,\xi))^2] \\
&= \min_{\beta,\gamma,\xi} E[(D - \hat{d})^2] \\
&= \min_{\beta,\gamma,\xi} [ED^2 - 2\hat{d}EX + \hat{d}^2] \\
&= -\, 2ED + 2\hat{d} \\
\text{so, } \hat{d} &= ED
\end{aligned}
$$

D is a vector, D = [D1, D2, D3], where each component is a metric from the data. D1 is average user duration. D2 is the scaled number of users becoming infected from between groups. D3 is the average number of new users per day. These three metrics together capture important aspects of the dynamics of the spread of the movement and hashtag. I combine the distance between these actual data metrics and their model based estimates, $m(\beta, \gamma, \xi) = \hat{d}$ into one error metric. I ensure all components contribute equally by dividing each component by its standard deviation and then add them together. I run each model over a series of possible parameter sets three times and take the average mean squared errors at given parameter sets.

The components of D differ for the movement and for the hashtag. For the movement I use average long user duration, the scaled total number who enter from between groups, and the average new users per day where a user can only be new once. For the hashtag I use average short user duration, the scaled total number of users who did not tweet the day before who tweet as a result of a between group connection, and the average number of new users such that the user did not tweet the day immediately before. The D from the data for the movement is D = [28.06, 15, 54.9]. The D from the data for the hashtag is D = [1.04, 26.5, 96.22].

The minimized distance between model output and the data for this combined metric is indicative of how well the model captures the data, but it is limited. This combined error metric made up of three pieces of information does not capture all of the intricacies of the phenomena both in the model and the data. It thus cannot be used as the only means of evaluation. I call the error minimizing parameters "optimal" in the sense that they minimize error, but it is possible that they are not the parameters that best match the model to the data. This is both because the metric is limited and because the true optimal parameters might exist outside of the parameter ranges I consider.

4.2 SIR

4.2.1 Popularity of #MeToo

We estimated from the data that $\beta \approx 0.14$ and $\gamma \in [0.0960, 0.0964]$ when modeling the popularity of the hashtag with an SIR framework. Below are graphs of populations over time and user durations from the data in the first row and from the model with data informed parameters in the second. Note that in the graph of populations over time from the data the recovered population is much larger than the infected population on all but the first day, so they are graphed on separate axes to remain visible. We see that the model predicted infection looks relatively short but the model user durations are still not as short as those in the data. The average user duration is 10.2 days. This is much longer than the average user duration of 1.04 days found in the data. This is reflected in the graphs of user durations. The infected population curve looks a lot like the first and biggest spike in the data for the number of users involved immediately following Alyssa Milano's tweet. Unfortunately, the SIR the model is unable to capture all of the following spikes in number of users tweeting #MeToo per day.

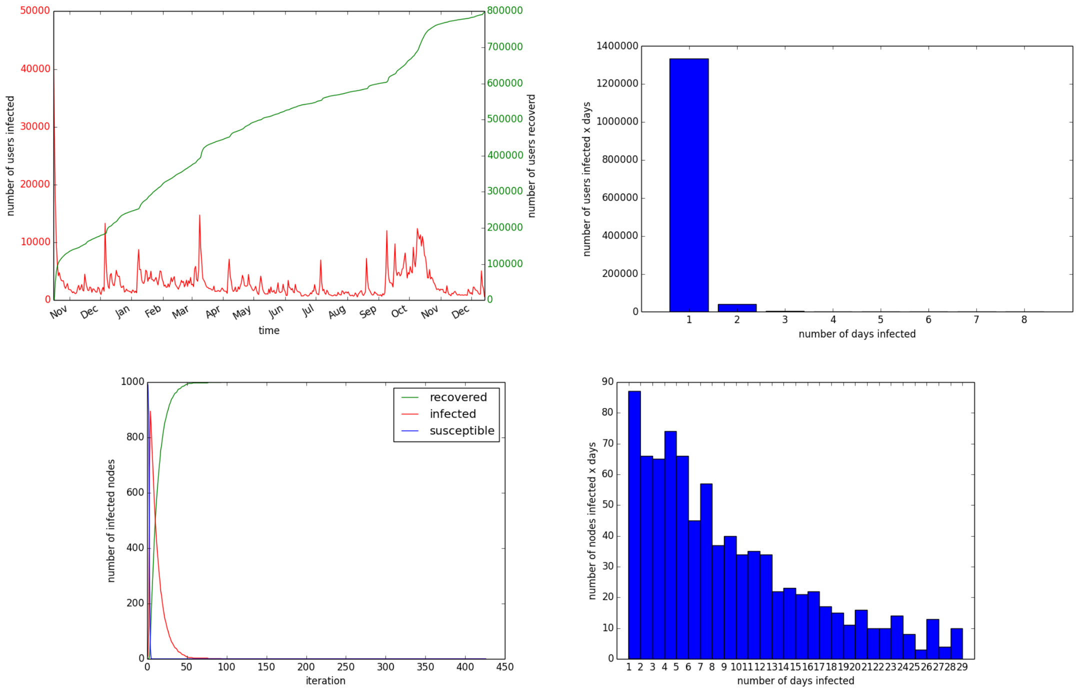


Figure 25: SIR populations over time and durations at β and γ estimates for modeling hashtag popularity in the data

Below is a heat map of the mean squared error for given parameter sets. The circle marks the approximate range for parameters as estimated from the data. The star marks the optimal parameter set that minimizes the distance between three key metrics from the data and those same metrics produced by the model simulation. The numbers on the axes are the fraction of 100 that result in a given γ and β. The optimal parameters that minimize the mean squared error with respect to metrics from the data are $\beta = .14, \gamma = .25$. With these parameters, the model has a minimum mean squared error of 0.43 with a standard deviation of 0.0481. The β estimated from the data is the same as the error minimizing β for an SIR model trying to capture the spread of the hashtag.

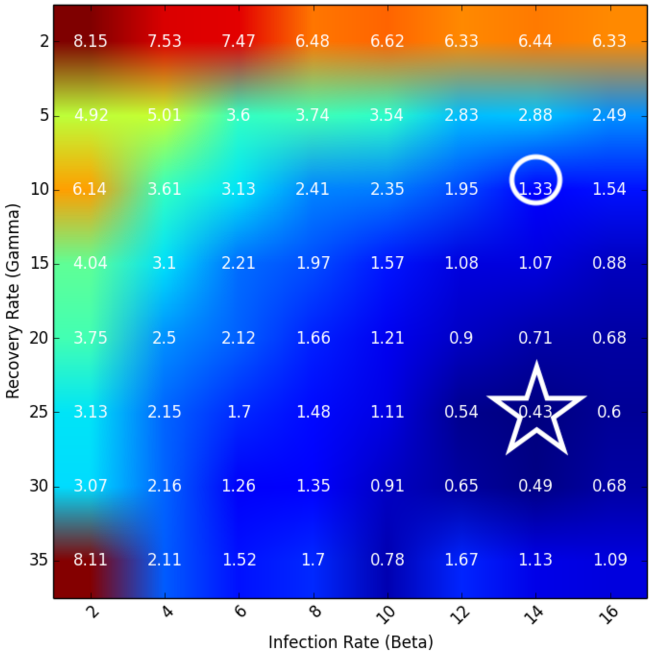


Figure 26: Heat map of error between SIR at given parameter sets and metric for the spread of #MeToo. Circle marks the range of estimated parameters from the data. Star denotes optimal parameters from the metric. Numbers on axes are fraction of 100.

I graph the populations over time and user durations from the optimal error minimizing parameters compared with the data. Using these optimal parameters we see one very quick spike in infection and then the infection rapidly dies out. The distribution of user infected durations is closer to the distribution from the data than with the previous parameters but is still not as extreme. The average user duration with these parameters, not surprisingly given the relation between γ and user duration, is 4 days. This is larger than the data average of 1.04 days but is closer than the parameter estimates.

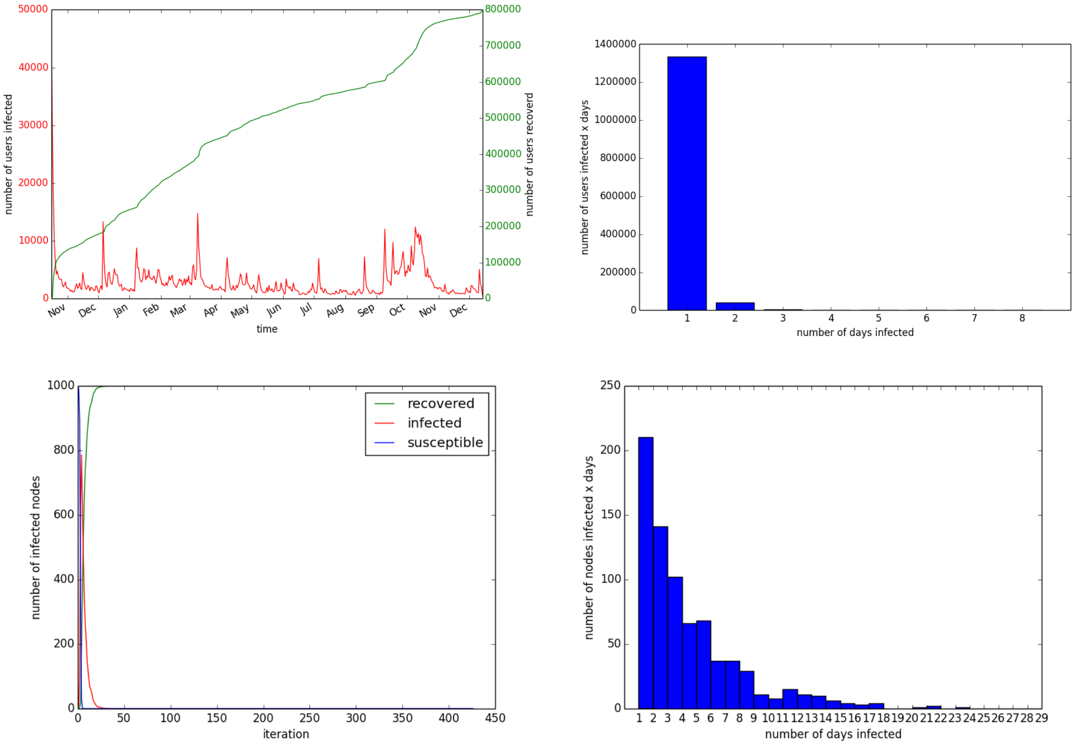


Figure 27: SIR populations over time and durations at β and γ that minimize mean squared error with respect to key data metrics for the popularity of the hashtag

An attempt to model the popularity of the hashtag with the SIR model can be interpreted really as modeling only the first big spike. A pure SIR model can only capture the first spike of the whole hashtag movement the best. Because all infected users become recovered and stay recovered, a second large spike in infection is not possible. In the SIR context, the day following Alyssa Milano's tweet was the spike and everything else is just noise at the end as the one bout of infection dies out. Because the number of users who tweet #MeToo is much greater on the first day than any other infection period, almost twice magnitude of even the next biggest spike, this is not an unreasonable assumption. An SIR framework provides a crude model for the popularity of a hashtag, but it has its limitations. For example, we know from the data that many users appear for more than one consecutive period of days, something that is not possible in SIR. There is no distinction between long and short infection in an SIR framework because each node is only possibly infected once. Though this distinction is clearly present in the data, the distinction is relevant only to the SISR and SIRS models in which nodes can be reinfected.

4.2.2 Me Too Movement

The estimates from the data for SIR parameters to model the movement as a whole are $\beta \in [0.00576, 0.00580]$ and $\gamma \in [0.035, 0.036]$. Below are graphs of populations over time and user durations from the data in the first row and then from the model with data informed parameters in the second. The recovered population in the graph over time from the data is still much larger than the infected population and so is plotted on a separate axis. We see that the shape of the infection curve in the model does not match the still infected population from the data. The infection here persists longer than it does with the SIR parameter set for modeling the hashtag, but it still does not cover the full 426 day period. I did not include the number of users who only tweet once in the user duration graph from the data because it dwarfs the rest of the distribution. The distribution of user durations from the data is skewed heavily right, even more so than may be captured by the graph from the data below. The graph of users by duration in the model does not exhibit the same magnitude of decay as the analogous graph from the data, but it is the same shape. The average user duration is 28.5 days. This is the same average user duration as the data suggests and is to be expected due to how we derived γ.

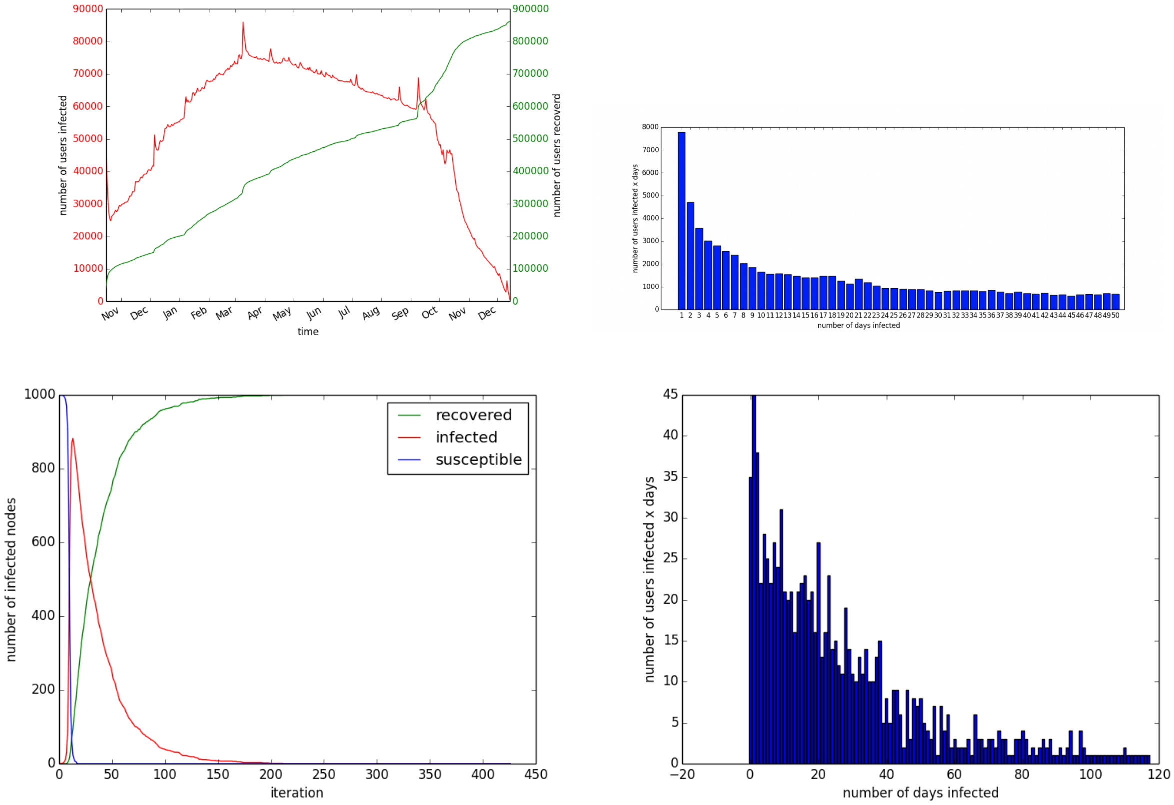


Figure 28: SIR populations over time and durations at β and γ estimates for modeling the movement as a whole

Below is a heat map of mean squared errors for given parameter sets. The minimum average mean squared error for a parameter set is 8.24 with a standard deviation of 0.946. This occurs when $\beta = 0.01, \gamma = 0.08$ as is labeled with a star on the heat map. The circle indicates the range of parameter estimates from the data. Both the mean squared error minimizing parameters and the estimates have a relatively low S to I transition probability, $\beta \leq 1\%$. A lower β slows down the infection rate and makes the SIR model's infected population over time look more like the still infected curve from the data. However, with a small β, sometimes the infection dies out before it can begin, with the first node or two infected becoming recovered before anyone else gets infected.

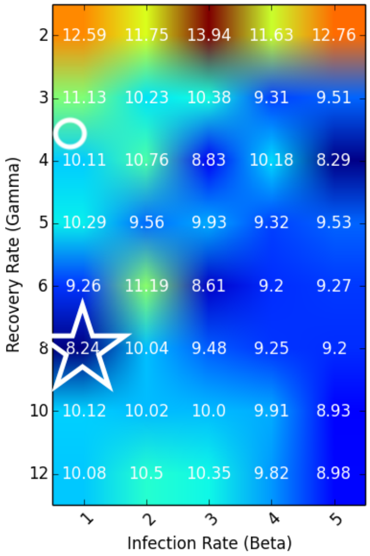


Figure 29: Heat map of error between SIR at given parameter sets and metric for the Me Too movement. Circle marks the range of estimated parameters from the data. Star denotes optimal parameters from the metric. Numbers on axes are a fraction of 100.

Choosing $\beta = 0.01, \gamma = 0.08$ creates the population over time and duration graphs shown below. The model infected population over time curve is much steeper and shorter in duration than the analogous curve from the data. Keeping in mind that the graph of distributions from the data below does not include users who show up for only one day, the distribution of durations under these parameters looks relatively similar to the data but still does not decay quickly enough. The average user duration from the model is 12.4 days which is much shorter than the 28 day duration from the data.

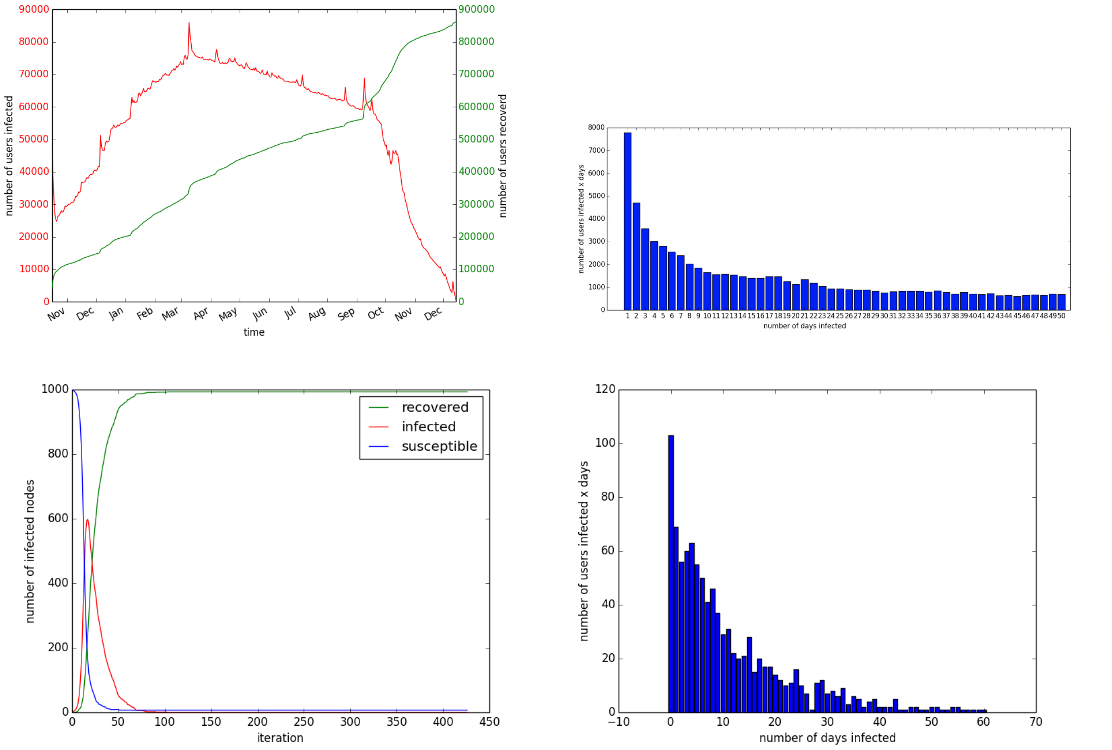


Figure 30: SIR populations over time and durations at β and γ that minimize mean squared error with respect to key data metrics for the movement

In theory an SIR model should be the ideal framework to capture the movement as a whole. This is because conceptually when someone begins to engage with the movement that person does not stop until the movement loses momentum and is no longer socially relevant. Once someone has engaged with the movement, it is hard for that individual to stop engaging as the movement continues to grow and reaches peak infection. Even if a supporter of the movement, for example, were to become a critic, that person would still be continuing to engage in the movement as a whole, just from a different perspective. Unfortunately, with both the estimated and optimally estimated parameters, the shape of the infected population curve and the corresponding user durations distribution is not quite right. Evidence suggests that an even lower β might result in a curve that better resembles that in the data. However, without a lower γ as well, the infection dies out quickly. And, if we decrease γ then the average user duration will be greater than 28 days. The inability of my simulation using an SIR model to capture a long sustained infected population with a relatively short user duration is likely a result of the size of the network over which my simulation is run. 1000 nodes is nowhere near the size of the network of 126 million daily active users on Twitter. On such a large network it may be possible to have a sustained infection with relatively short user duration as so many users tweet once, creating a large infected population on any given day, but not staying around for a long time to increase the average user duration. Because I am working with only a 10% sample, it is also possible, though not very likely, that the user duration distribution in the data in reality is not so heavily weighted toward users who show up only once and maybe the SIR model does a better job of capturing the true distribution.

4.3 SISR

Recall we do not have estimates for γ and ξ in the SISR model, just a relation. γ and ξ will satisfy the relation if $\xi = \frac{.37\gamma}{1+\gamma}$. I choose γ to be 0.10 and then ξ is approximately 0.03. I use this for γ and ξ here when modeling both the hashtag and the movement. Using larger choices of γ and ξ will produce similar population curves but a lower peak infection spike. The average short and long user distributions are also both shorter with larger γ and ξ. Using smaller probabilities for γ and ξ makes the peak infection greater and the average user duration longer.

4.3.1 Popularity of #MeToo

We estimated that $\beta \approx 0.14$ in the SISR model of the hashtag. The average user duration, that is short user duration, is 8.5 days. That is much longer than the 1.04 day average in the data. We also see that the distribution of user short durations does not decay any where near the rate that the distribution from the data does. The infection spikes and dies out like it does in an SIR model. Though in an SISR model, nodes have the possibility of reinfection, the infected population still dies out. The graphs of populations over time and user durations from the data on the popularity of the hashtag and from the model with estimated parameters are shown below.

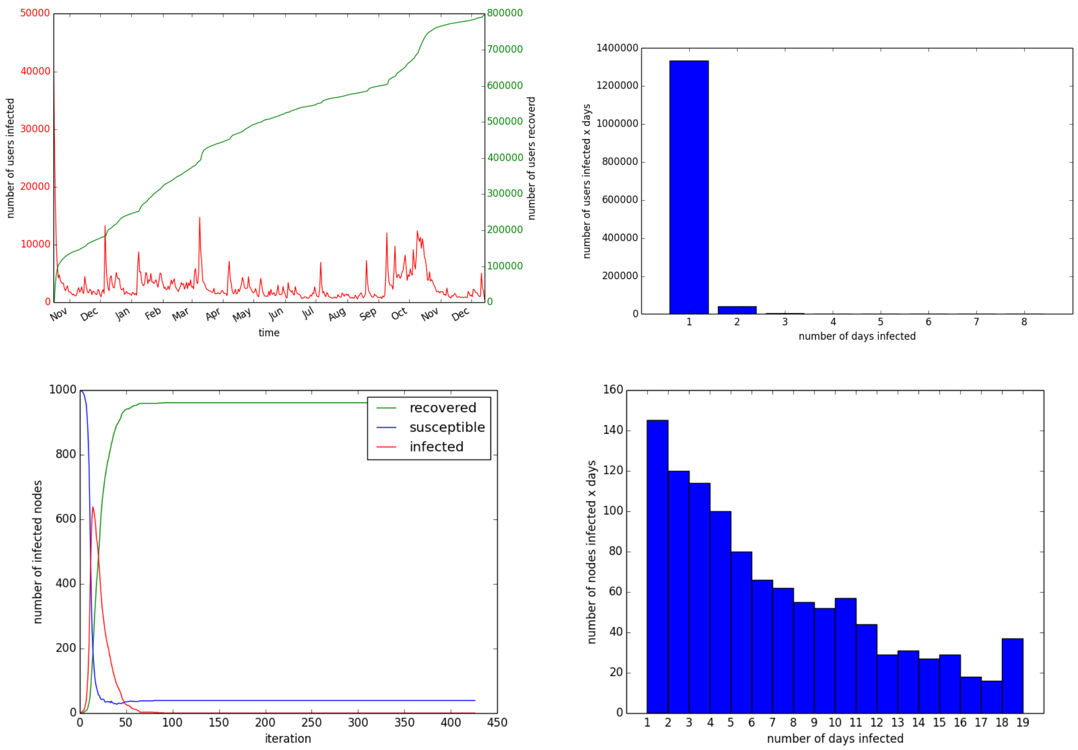


Figure 31: SISR populations over time and durations at β, γ, ξ estimated from the data for the popularity of the hashtag

Below are heat maps of the model error at estimating metrics from the data for given sets of parameters compared to the data for #MeToo. The optimal parameter sets are stared and the estimated ranges for β are circled. The axis numbers are fractions of 100. The minimum average mean squared error is 0.178 with a standard deviation of 0.0399 when $\beta = 0.08, \gamma = 0.30, \xi = 0.05$. These heat maps indicate that a higher γ and a lower β and ξ may be even more "optimal." When taken to an extreme, a SISR model with a high γ with a low ξ behaves like an SIR. A lower β means a longer time until infection.

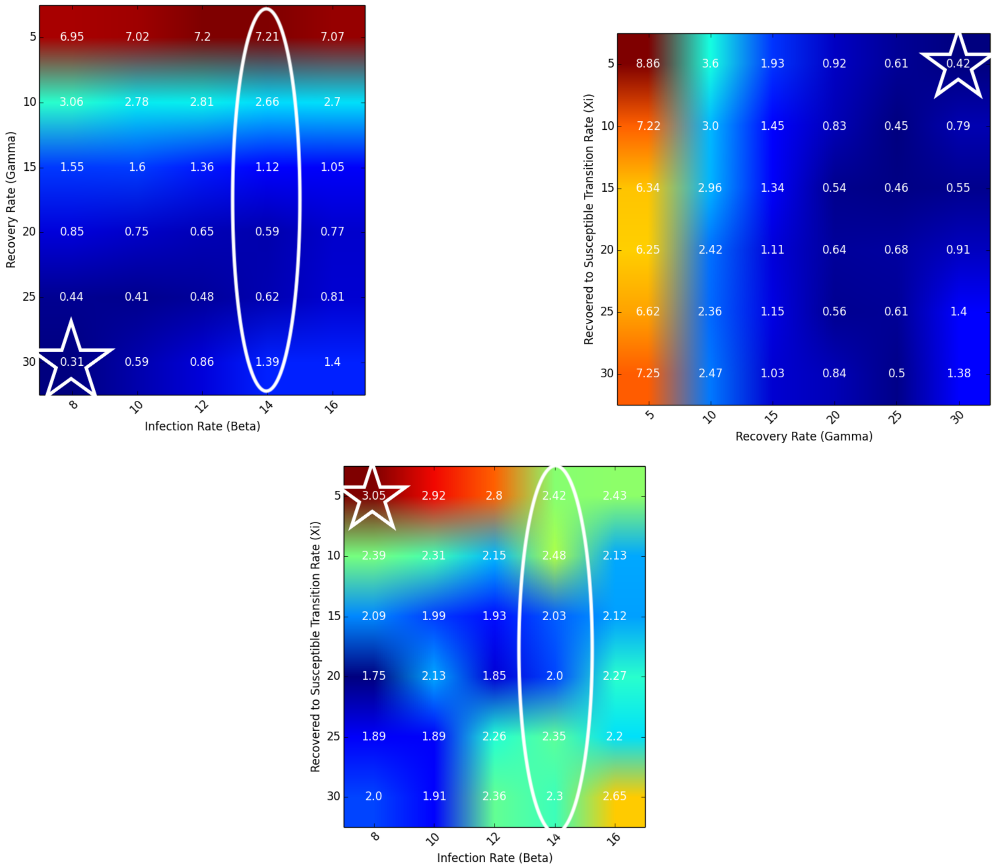


Figure 32: Heat maps of mean squared error of given parameter sets in an SISR model modeling the spread of #MeToo. Star denotes the optimal error minimizing parameters and the ranges circled are possible estimates from the data

Running the simulation with these optimal parameters does in fact result in population over time curves that look like a very short spiking SIR. The distribution of user durations looks more similar to the data than the distribution from the estimated parameters. The average user short duration is 2.3 days here. Below are the graphs from the data and from the model with optimal parameters respectively.

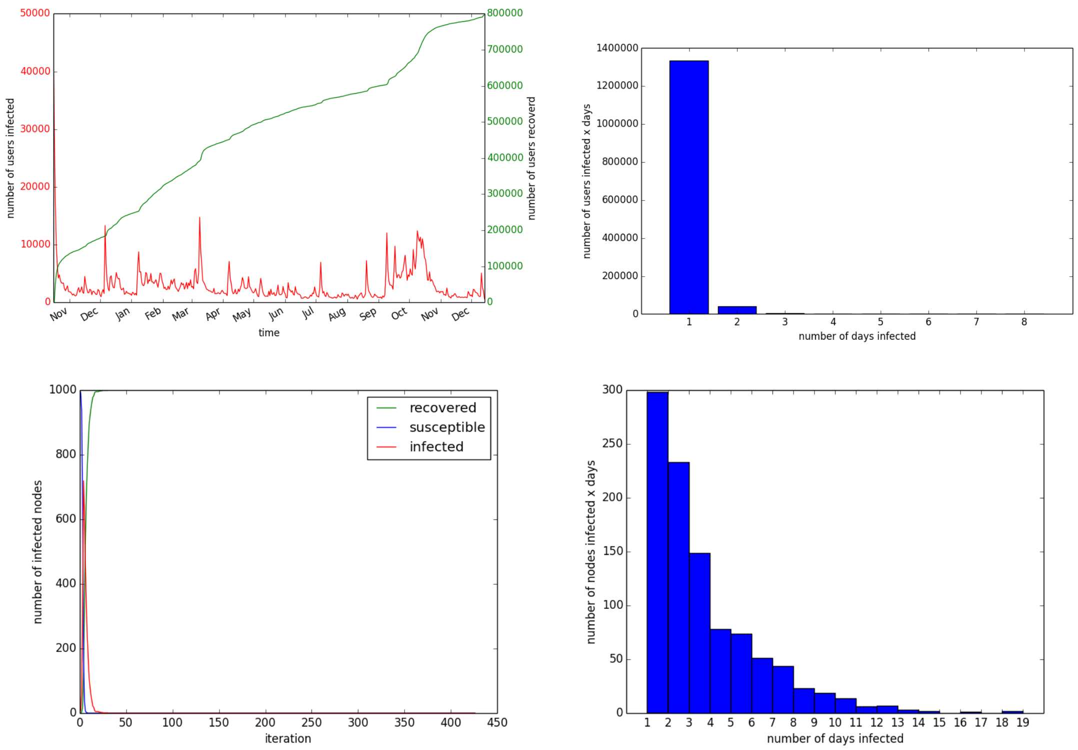


Figure 33: SISR populations over time and durations at the optimal error minimizing β, γ, ξ for the popularity of the hashtag

The fact that the more optimal parameters move in the direction of an SIR is perhaps a reflection of the inability of my error metric to capture intricacies of the data. The SISR with optimal parameters, similar to an SIR captures the first big spike in popularity of the hashtag. An SISR should, however, be able to do what an SIR can not and reinfect nodes, continuing the infection. In theory as some users in an SISR become recovered, they leave and the infection dies down, but because some other users are still getting reinfected, there continues to be smaller spikes in infection. This is the exact activity that we see in the data for the popularity of the hashtag. The fact that neither the parameter estimates nor the optimal parameters provoke this behavior in the SISR is again perhaps attributed to the small, densely connected network over which my simulation is run.

4.3.2 Me Too Movement

For an SISR model attempting to capture the movement we approximated $\beta \in [0.00576, 0.00580]$ with γ and ξ related as they were before. Below are the graphs from the data of populations over time and user durations compared with the same graphs from the model. Because SISR has nodes that can get reinfected, here I graph the curve of users that are still infected. The average user duration is 9.8 days, much shorter than the 28 day average from the data. The distribution of user durations has a reasonable decay, but appears too heavily weighted around short distributions to achieve an average duration of 28 days. The small β extends the time until infection to be more similar to the curve of still infected from the data, but with a smaller β, the infection struggles to continue. In fact, in some iterations of the simulation with this model, the infection did not spread beyond one node before that node was recovered.

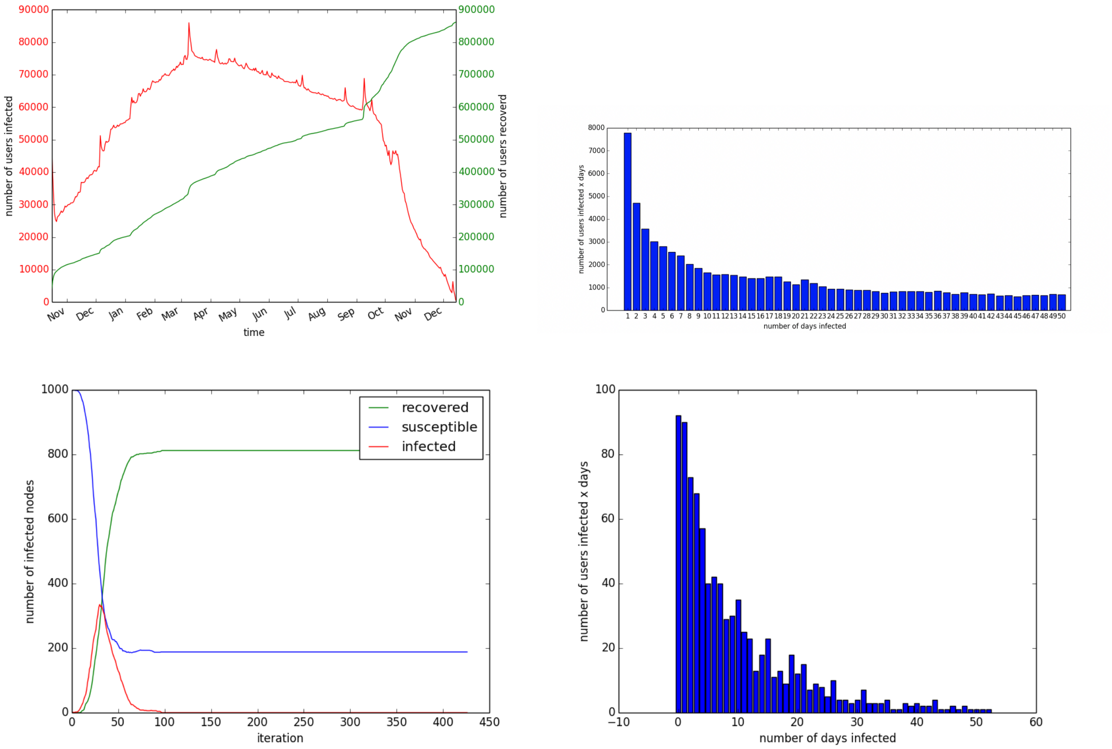


Figure 34: SISR populations over time and durations at β, γ, ξ estimated from the data to represent the movement

Plotting these parameter estimates on error minimizing heat maps shows that at least the estimate for β is not far off from the error minimizing optimal β. The minimum average mean squared error between the metrics from the model and from the data is 1.244 with a standard deviation of 0.0824 when $\beta = 0.005, \gamma = 0.05, \xi = 0.3$. Note that the numbers on the axes below are out of 1000.

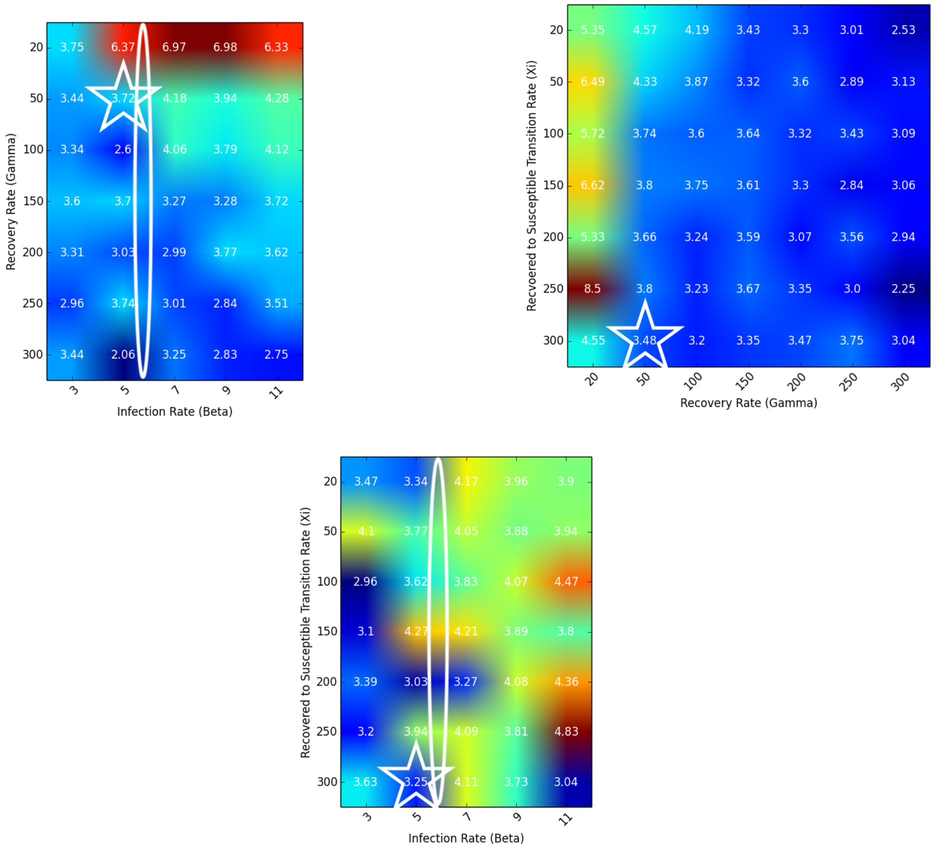


Figure 35: Heat maps of mean squared error of given parameter sets in an SISR model modeling the movement. The star denotes the optimal error minimizing parameters and the ranges circled are possible estimates from the data. The numbers on the axes are out of 1000.

I plot the curves for the movement from the data below and compare to curves generated with the optimal parameters. The average user duration with the optimal parameters is 3 days. But because there are not many nodes that get infected, there is no real distribution to compare. Because the optimal β is so small, the infection never sees a rapid viral spike. In fact the infection only ever reaches a maximum of 5 nodes. However, the slow, non-viral growth of the still infected population looks more similar in shape to the analogous curve from the data than anything else we have seen.

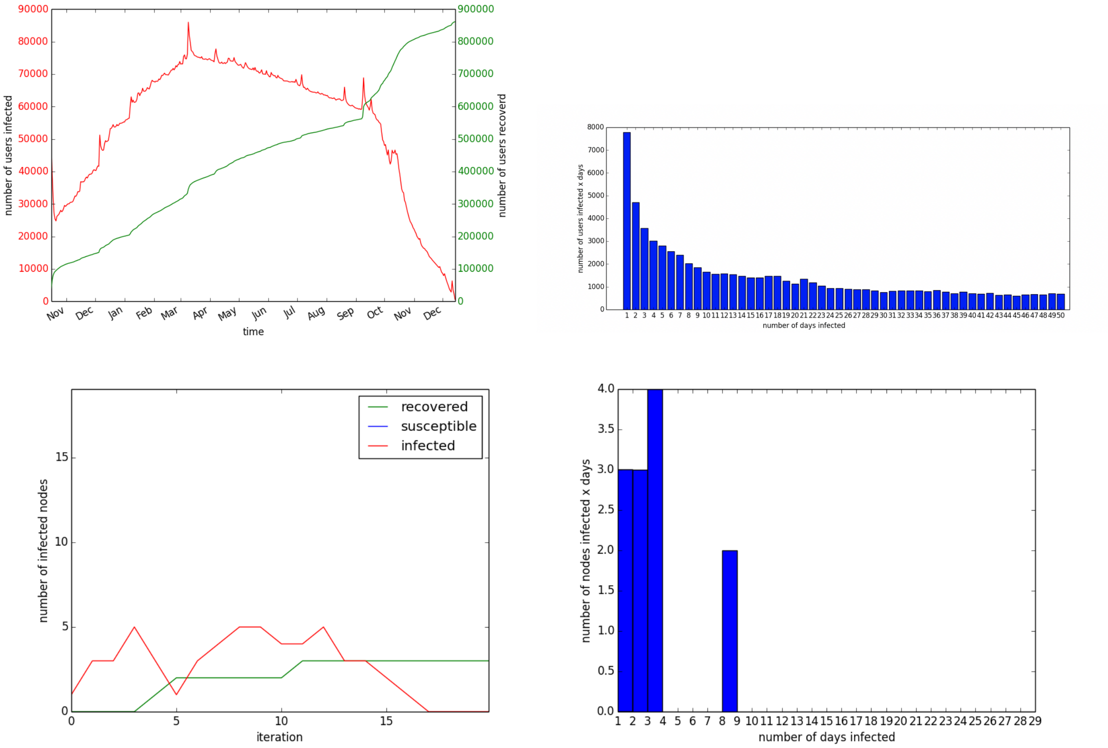


Figure 36: SISR populations over time and durations at the optimal error minimizing β, γ, ξ for the movement.

The fact that slow growing shape of the still infected population from the model looks somewhat similar to same curve from the data reinforces the idea that β must be small to get such a shape. However, with β small, it is difficult for infection to spread before dying out. The infection in the simulation with the optimal parameters above reaches 12 nodes before dying out. One could argue that perhaps if run on a larger network, the still infected population would be larger. But we see in this simulation that 12 nodes or 1.2% of the population is infected and in the Me Too movement 3.3% is infected. So, the inability to perfectly capture the movement as a whole may be a structural of the underlying network or of using SISR as a framework.

4.4 SIRS

Attempting to capture the popularity of the hashtag, we estimated from the data that $\beta \approx 0.14, \gamma \in [0.0960, 0.0964], \xi \in [0.0079, 0.0082]$. The average short user duration is 1.04. This is spot on to what it was in the data as is to be expected because of how we derived γ and because people in the I population transition only to R. The distribution of user durations looks impressively similar to the distribution of user durations from the data. And, the graph of populations over time below shows a relatively big spike in popularity in the beginning and then a sustained level of background noise with small peaks in activity. Despite not having enough larger spikes, SIRS visually appears to capture both the general shape of populations over time and the user durations.

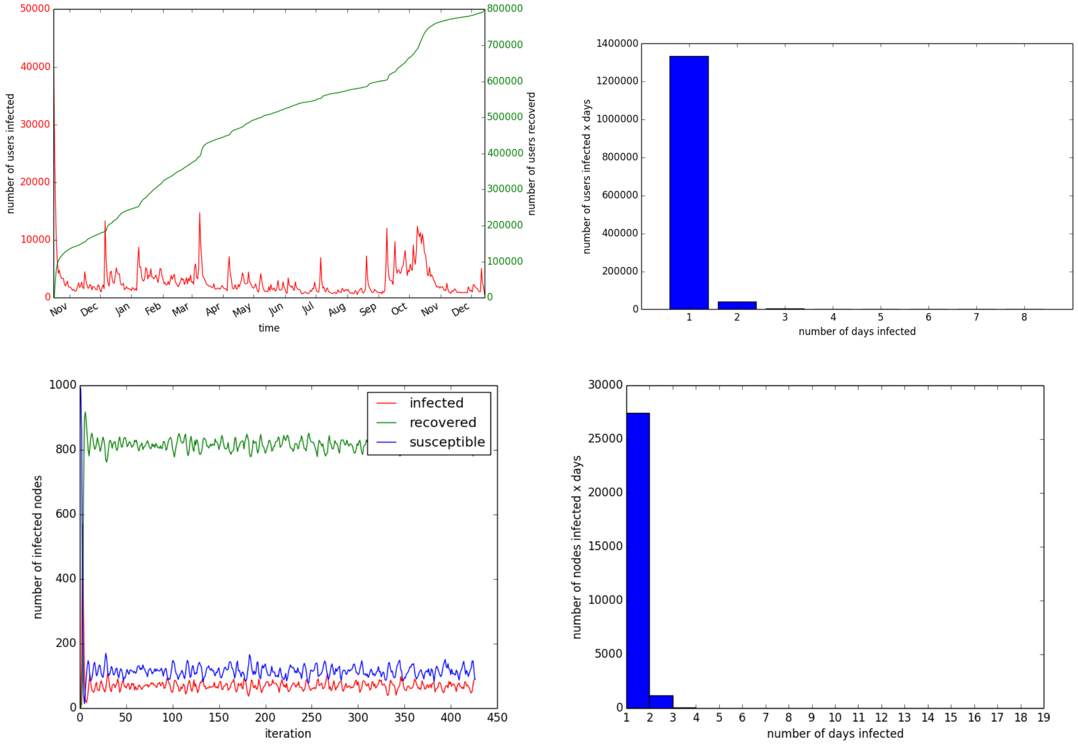


Figure 37: SIRS populations over time and durations at β, γ, ξ estimated from the data for the popularity of the hashtag

I plot heat maps of the error of the model output compared to metrics from the data. The optimal parameters are stared and the estimated parameters from the data are circled. The numbers on the axis for γ and ξ are to fractions out of 1000 and for β are fractions out of 100. The minimum average mean squared error is 18.32 with a standard deviation of 0.02738 when $\beta = 25, \gamma = 120$, and $\epsilon = 7$. The heat maps indicate that a higher β and γ and a lower ξ might be more optimal. Taken to an extreme this might cause the model to behave more like an SIR, or at least slow the time it takes to transition from the recovered to the susceptible population.

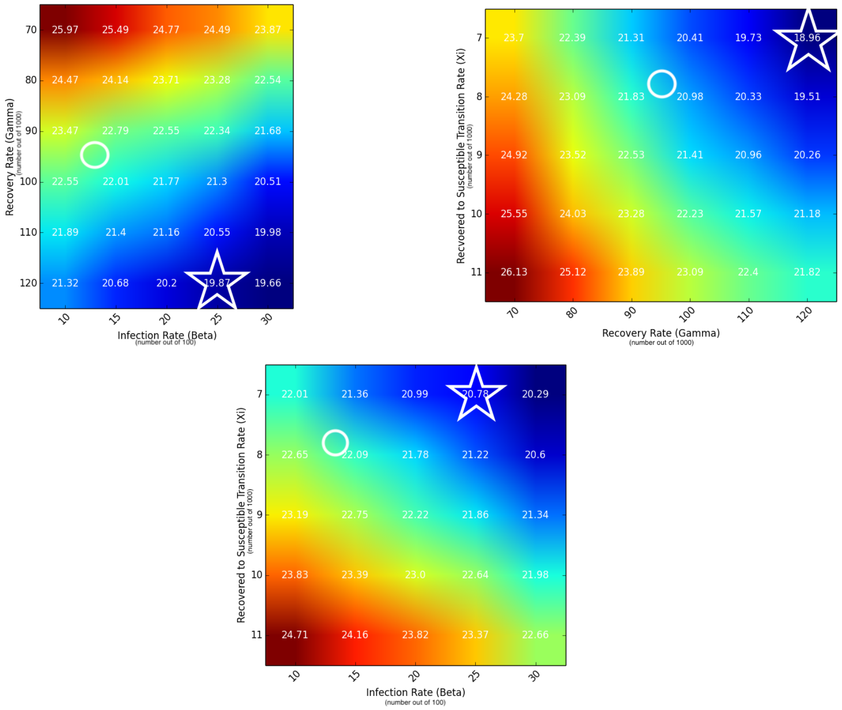


Figure 38: Heat maps of mean squared error based off of a metric for an SIRS with given parameters β, γ, ξ. Star is optimal set of parameters and circle is the range of parameters suggested from the data

I plot the optimal parameters and compare to the analogous plots from the data below. The average short user duration is 1.0 day. With β small and $\beta \leq \gamma$, many iterations saw the infection immediately die out either with the first node or the first few. When the infection caught on it looked like the following graph of populations over time in the second row, displaying the cyclical behavior that is known of an SIRS. All iterations in which the infection caught on saw an initial spike that was delayed relative to the data and smaller. The distribution of user durations shows us that in the model all users show up for a period of one day at a time. This is even more extreme than the distribution from the data.

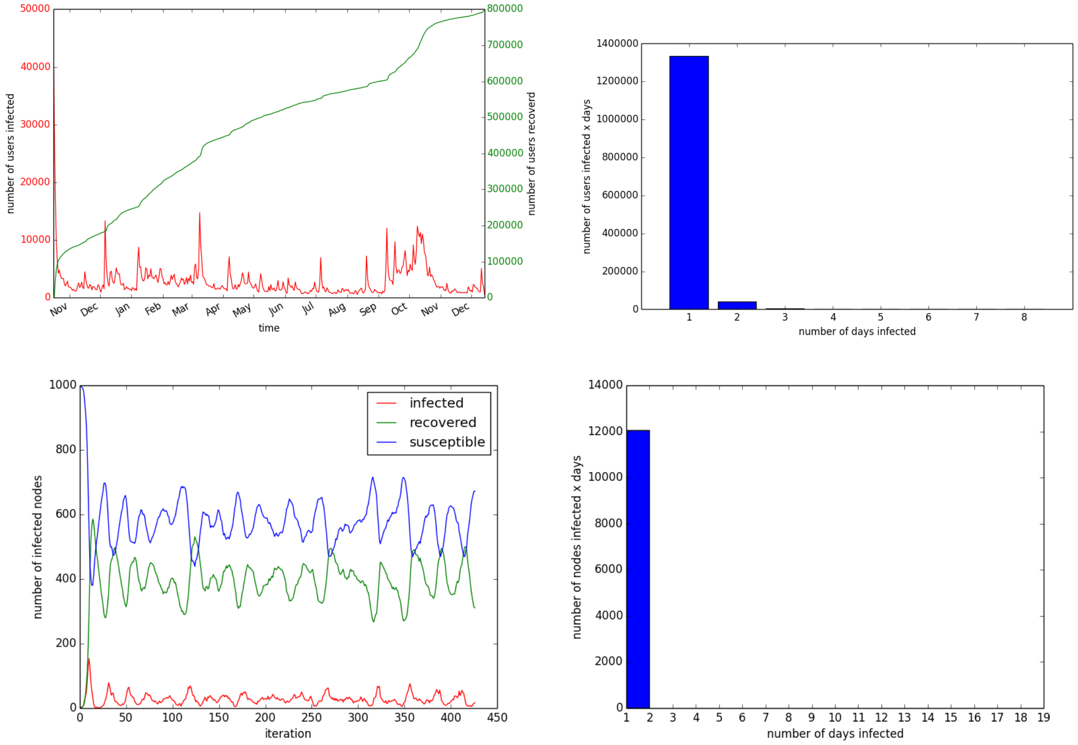


Figure 39: SIRS populations over time and durations optimal error minimizing β, γ, ξ for the hashtag.

The SIRS model seems to capture the popularity of the hashtag very well, qualitatively better than the other models. However, even at its optimal parameters, it still has the highest minimum mean squared error for a model of the hash tag popularity by far. Let us consider the metric for mean squared error. In the case of SIRS, two of the three components of the error metric seemingly contradict themselves, average new users per day and average user duration. In the SIRS model there are only new users at the very beginning and then users get reinfected. If only a small number of users ever get infected but then keep recovering and reinfecting each other, the metric for average short user duration will be spot on, but there are hardly any new users. The same is true if all users in the network get infected early on and then just keep reinfecting. Either way, this could cause the error to sky rocket. With a small group of users reinfecting each other, not only would the average number of new users per day be off, but the number of connections coming from between groups would be off as infection may occur locally. Thus we see a possible flaw of using the SIRS model to capture the dynamics of even the spread of the popularity of the hashtag.

Recall that parameters to model the Me Too movement with the SIRS model were not logical to pull from the data. The SIRS framework does not structurally make sense to model the broader overarching movement because of its strong cyclical patterns. To see this I have plotted the long user infection and recovery curves using the parameters estimated for modeling the hashtag below. If SIRS were to model the movement it would look like the following. The user durations are completely skewed in the wrong direction and the infected and recovered curves are basically meaningless, signally only that the simulation iterations stop running near the end. The average long infection duration is 409.5 but the average short infection duration is 1. Nodes will be infected for only one day at a time but stay in the total population forever because they can not fully recover. Heat maps of SIRS for the movement and the corresponding optimal parameters can be found in the appendix.

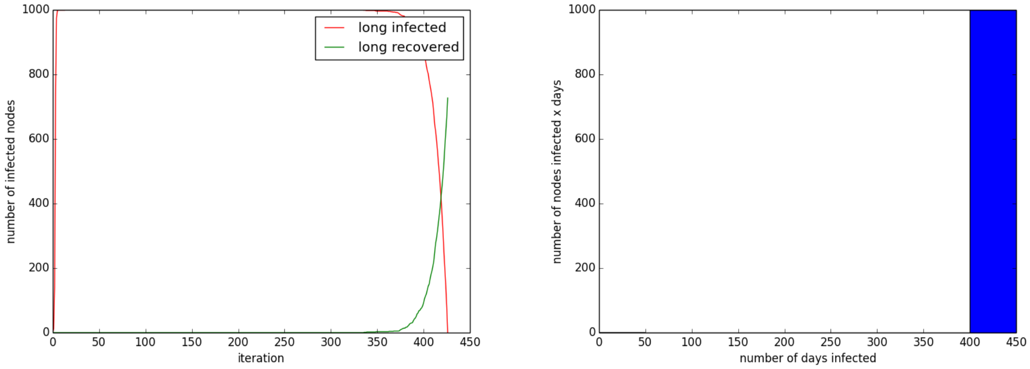


Figure 40: Populations over time and user duration for modeling the movement with SIRS and parameter estimates from data for the hashtag

An SIRS model is plausible to think about in the immediate short run as long as the movement remains socially relevant. Users, even if they are not tweeting, are likely still engaging with the movement to this day even just by reading the news. Thus, as of right now it might be reasonable to assume that the broader society as a whole never recovers. However, based on previous social movements, we know that the infection generally dies out after some longer period, something that can be captured by a model framework like SIR or SIRS.

An SIRS model may also logically be applied to the extreme long run because people will likely never stop caring about sexual assault, woman's rights, and related causes. In this extreme long run frame of mind, the Me Too movement would be just one short period of infection in a broader cycle of activism and inaction around the related topics.

4.5 Between Group Connections

One of the pieces of information used in the error metric was the scaled number of nodes entering the infected population from between group connections. A graph of populations over time in an SIR simulation with data estimated parameters for the movement is pictured below on the left with each group's infection plotted separately. If β is the same for both groups 1 and 2, where group 2 is the smaller group with about 20% of the network, then the infections look approximately the same, just smaller in group 2 than in group 1. If β is much larger in one group than the other, then peak infection for both groups happens sooner and is larger, but there is some difference in the shape of the infected curves.

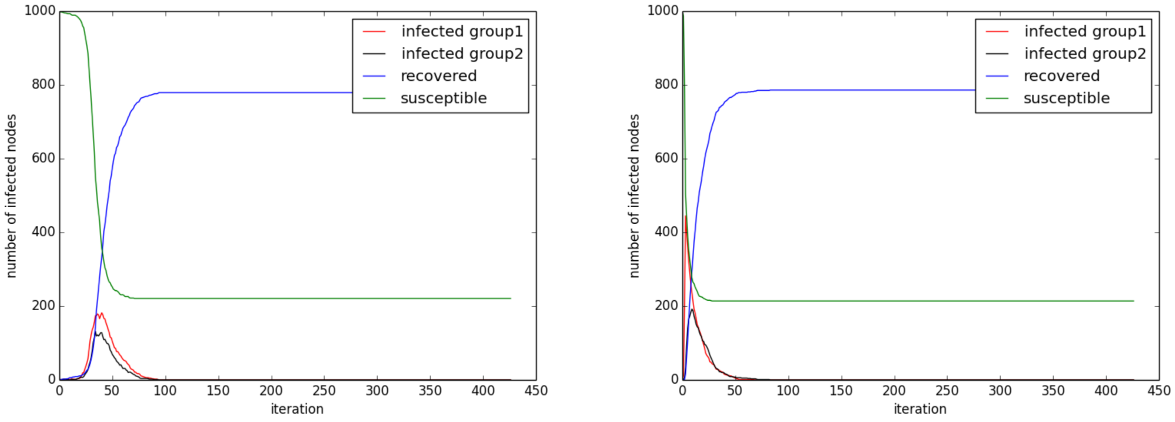


Figure 41: Graphs of SIR infection with data estimated parameters broken down by group. On the left the β's for each group are the same and on the right, β for group 2 ¡¡ β group 1

The question of cross group dynamics remains interesting in modeling social movements. There is opportunity for more analysis to be done on varying all parameters of the models between groups. Another way to evaluate infection of the Me Too Movement is to look at the amount of Google searches in different countries. Google trends data show that searches for the movement and the hashtag in the United States have remained at a relatively higher sustained level of interest as a fraction of its peak popularity than it has in other countries [1]. Other countries might see a high level of activity for a period of time, but these bursts of activity are not sustained the way the infection is in the U.S. This is perhaps a result of the network structure allowing users in the U.S. to keep reinfecting each other. While infection spreads internationally, the groups in other, especially smaller or less developed, nations do not reach levels of infection that allow them to keep reinfecting each other in a sustained way. This is also reflected by the fact that the ratio of international users retweeting U.S. domestic users to international users retweeting each other is 0.45, much higher than 0.05, the same ratio for the U.S. The U.S. functions as a highly infectious community, sustaining a long period of infection and occasionally spreading it to other international communities. It would be interesting to further explore changing β for communities in the network in an attempt to capture this dynamic.

The number of connections being made between groups is highly dependent on network structure. Recall that the 1000 node network that I graph all simulations over had a .1 probability of connecting within group and a .0125 probability of connecting between groups when it was constructed. Holding constant the probability of connecting within group at 0.1, below is a graph of the number of nodes becoming infected from a between group connection. As is to be expected, as the network connections between groups in a network increase, so does the number of nodes getting infected from a cross-group connection.

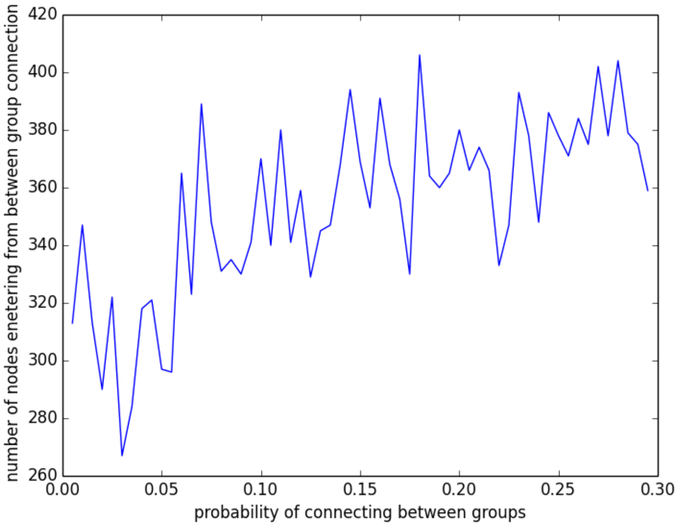


Figure 42: Plot of number of nodes entering from between group connection given probability the graph connects between groups for a basic SIR with parameters estimated from the data

Analysis done on examining how infection differs between groups with different network structures revealed networks with multiple tight communities and some connections between communities are best at spreading information [23, 25, 33]. It would be interesting to further explore the two broad international and domestic groups on a more refined scale and look at how tightly connected each country's network is. It would be especially worth while to do this in conjunction with finding different β's for each group to see if, with a set universal β, the connectivity of communities alone can account for their infectiousness.

4.6 Reintroduction of Exogenous Infection

As discussed in the data section, it is very possible that the popularity of #MeToo is influenced by exogenous news sources. I attempt to capture this with introducing reinfection into my models. Every 25 iterations I reinject infection into 1% of the network in an attempt to capture, say, a big news story breaking. 1% was the average amount of people retweeting mainstream news outlets and is thus perhaps a proxy for the amount of people tweeting in response to the news. Below is a graph of populations over time in an SISR model with parameters estimated from the data for the hashtag with reinfection. Because the reinfection is small, this simply adds a small amount of noise to the decay of the big spike and draws out the infection longer.

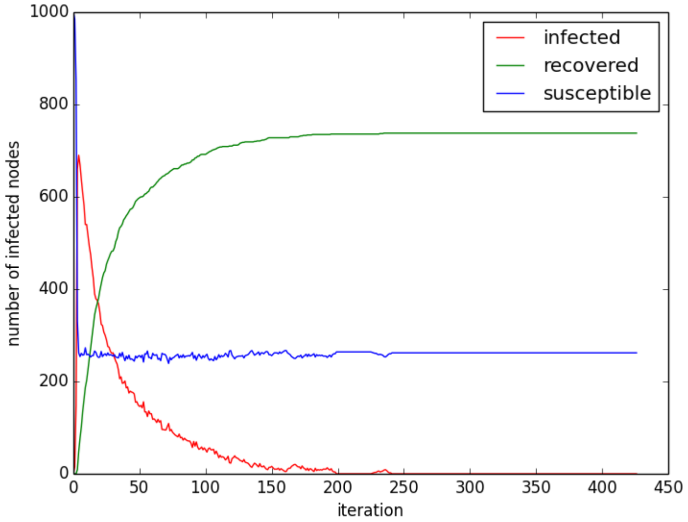


Figure 43: SISR with data estimated parameters for the hashtag and reinjection every 25 iterations

Though not shown here, plots look similar with even less added noise as a result of the reinfection for SIR because of the short duration of the spike.

Because of the cyclical and noisy nature of SIRS populations, reinjecting infection into the network is irrelevant in that framework because the nodes become reinfected themselves anyway.

Further work could be done here to better approximate the amount of influence exogenous to the network on users tweeting #MeToo and how important that is to the sustained level of infection.

5 Discussion and Conclusion

From the above analysis we see that the SISR model minimizes mean squared error for both the hashtag and the movement as a whole, making it the best model framework out of the three tested at capturing social movements, particularly the Me Too movement. SIR works to model the first big spike of a hashtag movement and in theory can capture the dynamics of the movement as a whole. SIRS makes sense as a model framework only to capture the popularity of the hashtag. It does so qualitatively well despite a large mean squared error estimate. Unfortunately, all frameworks and modeling attempts have their limitations. The analysis done in this paper is subject to a variety of constraints in the data, the network, and the model frameworks.

First, the data set is limiting in two ways. One limitation of the data is that it is only a 10% cross-sectional sample and thus we are not able to capture all of each user's activity. We may therefore have more users who show up once than is true in the full data, and our estimates about user durations could be subject to error. The second data limitation is that the data set only contain tweets with #MeToo in them. Ideally, I would be able to reconstruct the actual Twitter network and see where and when #MeToo shows up. This is what the model and simulation do, but due to the #Metoo limitation I must analyze only proxies for this in the data. Also, having data only about the hashtag is constraining when thinking about the Me Too movement. Though the name was popularized starting with Alyssa Milano's tweet, the modern day #MeToo movement quickly expanded beyond twitter. Thus, it is unlikely that a model based on Twitter data will fully be able to capture the dynamics of the Me Too movement as a whole.

The network over which I ran my simulations is a big limiting factor in the ability of my model to best capture #MeToo and the corresponding movement. The simulations I ran frequently resulted in near almost 100% of the nodes in the network being infected at some point. The high saturation of infection in my network is possibly indicative that the estimates for β were too high, but is more likely the result of the size and structure of the underlying network. When constructed, the network had a 10% probability of connecting with a node in the same group. This would imply that each user is likely connected to 10% of the users in the same group. This is a pretty extreme over estimate of in group connectivity. For example because there are 27 million DAUs in the U.S. this would imply that each user in the U.S. followed and had at least 2.7 million followers, an absurdly high amount. Thus, the network over which I ran my simulations is too connected to be very realistic. The extreme connectivity allows for higher and quicker infection saturation. The small size of the network is also not very realistic. I had to keep the network at 1000 nodes due to computing restraints, but a 1000 node network is on an order of magnitude of 10^5 smaller than the actual Twitter network. Such a small network is not able to accommodate a large percentage of users tweeting only one day while maintaining a sustained infection over the duration of the 426 days. Also, when the small network becomes so saturated with infection, it is difficult for new users to enter at the rates they do in the data. The fact that bootstrapped parameter estimates from the data had such low error ranges, gives us confidence in those estimates. The fact that those parameters did not always create results that captured the spread of the hashtag or the movement well means that there was something off about the underlying network, the model framework, or both.

Finally, as discussed throughout the evaluation section, the three different model frameworks are each constraining in their own ways. The SIR is unable to capture any sort of reinfection, something that is crucial to accurately model the popularity of the hashtag. Conceptually, SIR is a good framework to model the movement as a whole. The fact that its minimum mean squared error was relatively high (8.4) may be a reflection of significant constraints of the underlying network or a failure of the error metric to capture the important aspects of the movement. SIRS by far had the highest mean squared error in modeling both the hashtag and the movement. Though SIRS does not conceptually make sense for modeling the movement as a whole, it does for modeling the hashtag in the short run. In fact, it is the best at recreating infections with short user durations and many small infection spikes that both resemble what is seen in the data. The high error is likely a result of network saturation and the resulting inability to capture new users. Finally, SISR as a framework makes the most conceptual sense in attempting to capture the popularity of the hashtag. It also logically fits the pattern of user engagement with the movement as a whole. SISR has the lowest minimum mean squared errors for modeling both the hashtag and the movement. The only problem in both cases is that the model predicted sustained infection is relative short and user durations are relatively long. For reasons discussed above, these problems are most likely largely explained by the small, densely connected network on which the simulation is run. There may also be error in the way I constructed my simulation to run SISR. I first see if an infected node will become recovered, then if it does not I see if it becomes susceptible again with two independent probabilities. This means the probability of transitioning to S from I in my simulation is really $(1-\gamma)\xi$, making it more likely nodes will recover rather than get reinfected, and the total infection will die out sooner.

Overall, the SISR model framework is the most promising for future social and political movements online. There is exciting opportunity for further research in two related areas either applied to the Me Too movement in particular or social and political movements more broadly. First, it would be interesting to explore how much of an effect exogenous influences have on the conversation and levels of infection online. Second, a more in depth analysis of between group connectivity and infectiousness would be relevant. The data suggests that #MeToo is significantly U.S. dominated and the infection in the U.S. community will infect others internationally more strongly than the reverse. It would be interesting to attempt to capture that dynamic with varying β's and multi-community network structures representative of the international online landscape. A further extension is to construct groups and communities ideologically, not just geographically as data suggests this is also a relevant division in the network. This type of analysis both geographically and ideologically is important to help us understand how movements and ideas spread online.

Extremism has taken hold in certain communities and is growing and spreading online. The shooter who recently shot and killed 50 and injured another 50 people in a mosque in Christchurch, New Zealand cited various far right influential individuals online that helped solidify his radicalization as a white-supremacist in a 74 page manifesto before he committed the act [15]. Various individuals mentioned in the manifesto including Donald Trump and Candace Owens show up as influential in the #MeToo Twitter data as well. Further research and solidification of a good model of infection spreading between ideological and geographic communities may help us capture and estimate the reach of such polarizing and dangerous content online and allow us to prevent future catastrophes.

6 Appendix

Figure 44: Oct 16th, 2017 Retweet Network

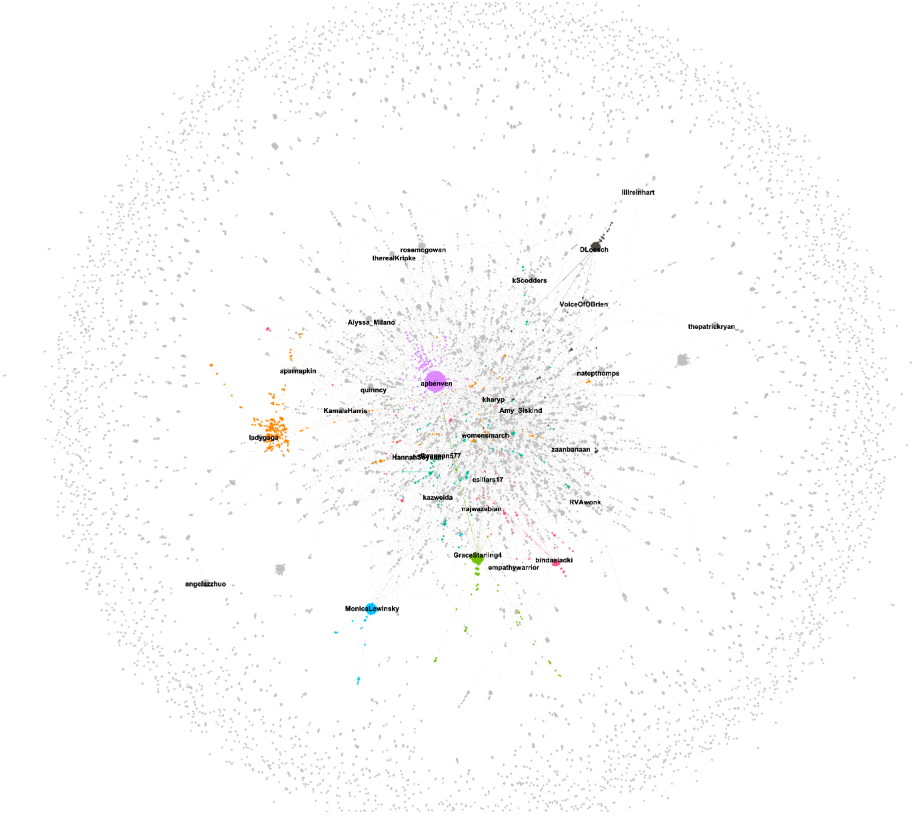


Figure 45: December 6th, 2017 Retweet Network

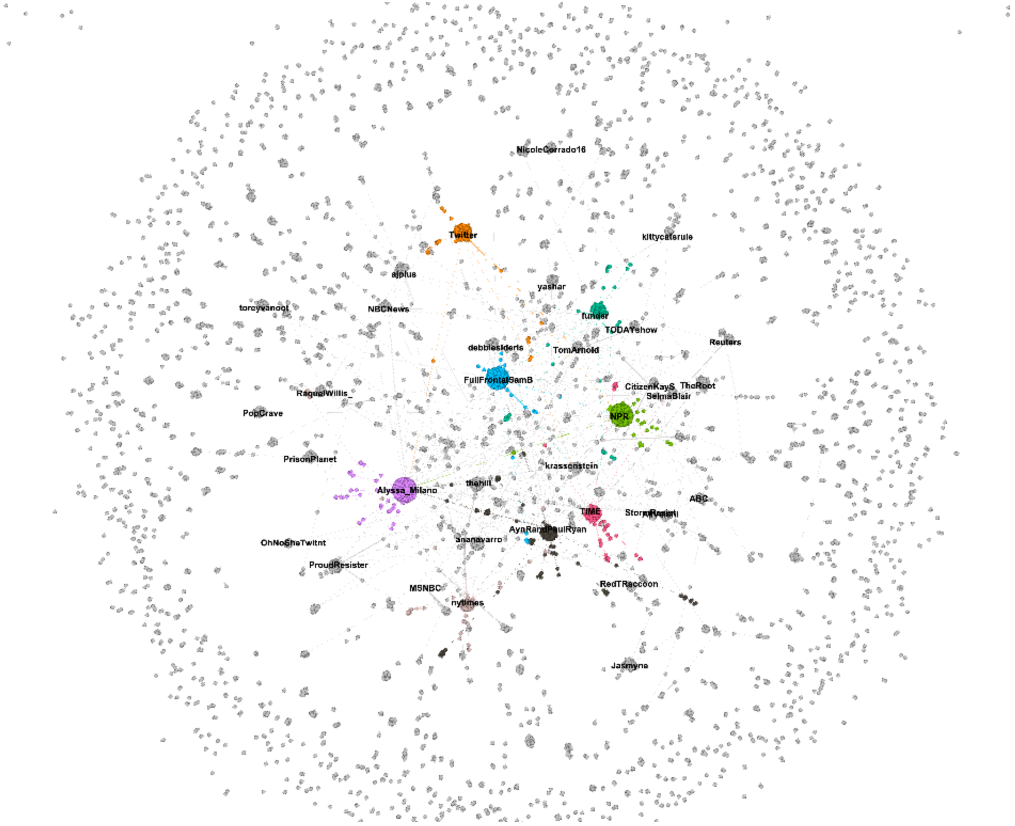


Figure 46: Dec 17th, 2017 Retweet Network

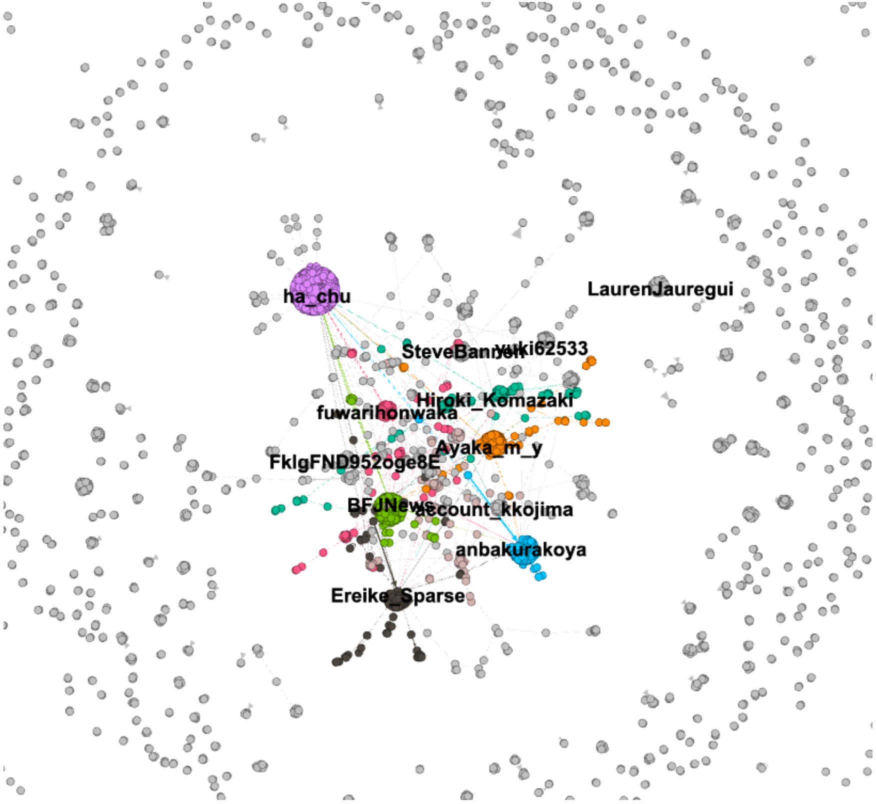


Figure 47: Jan 8th, 2018 Retweet Network

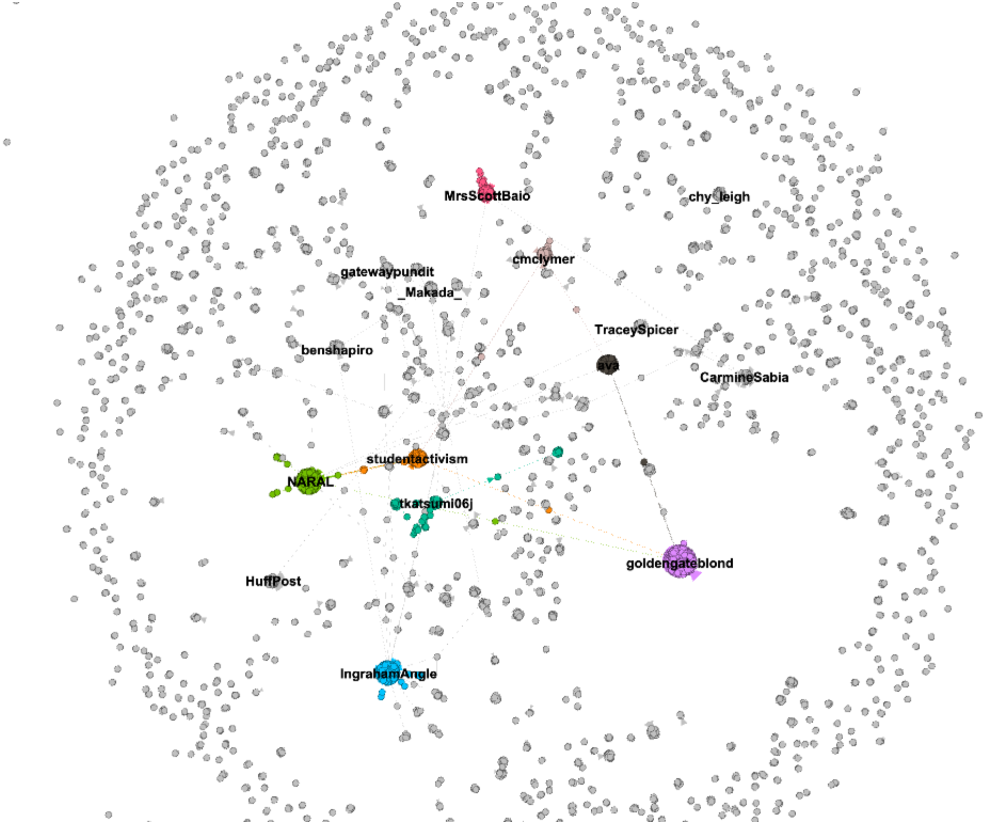


Figure 48: January 15th, 2018 Retweet Network

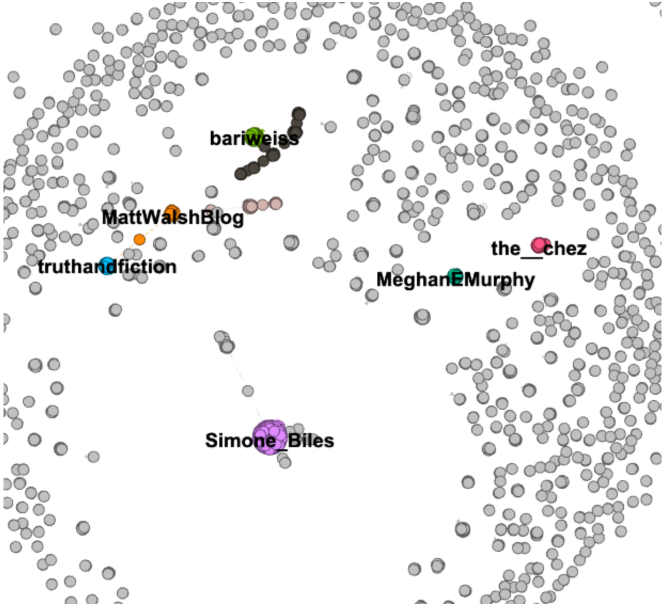


Figure 49: Jan 20th, 2018 Retweet Network

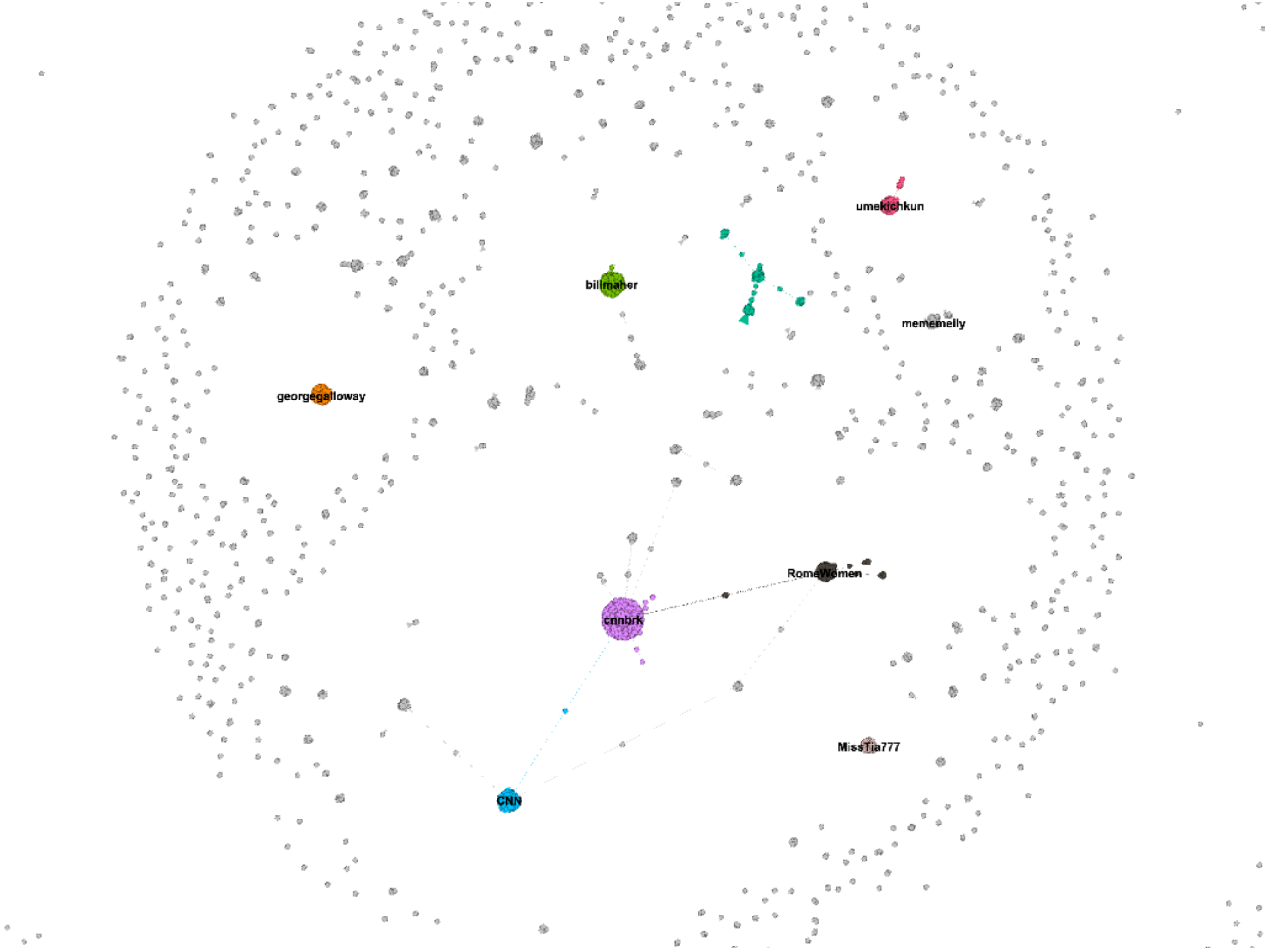


Figure 50: January 29th, 2018 Retweet Network

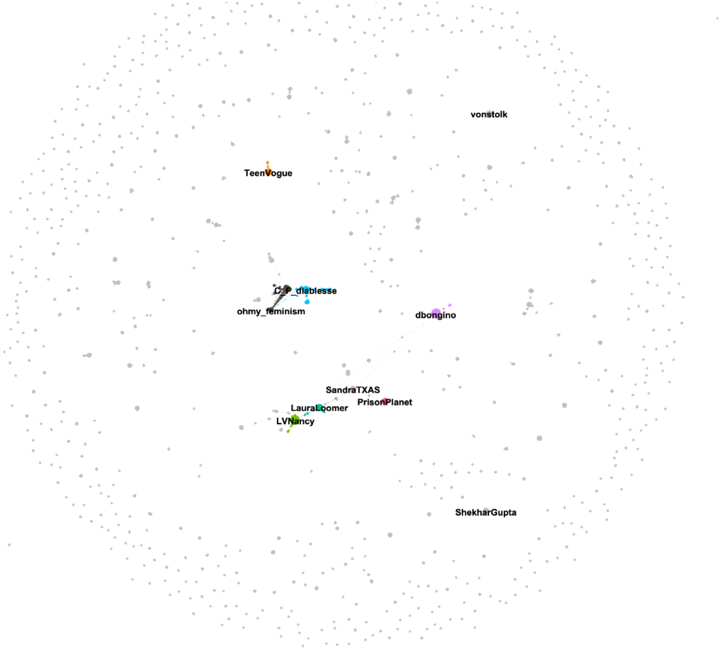


Figure 51: March 9th, 2018 Retweet Network

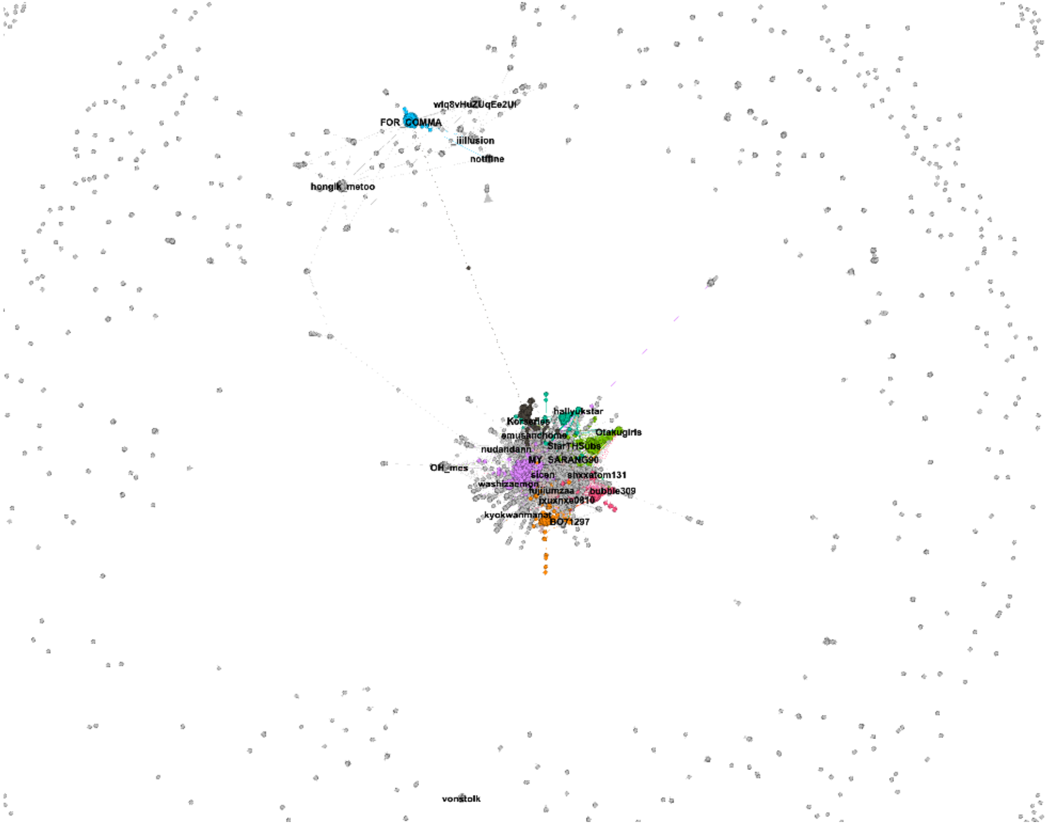


Figure 52: April 7th, 2018 Retweet Network

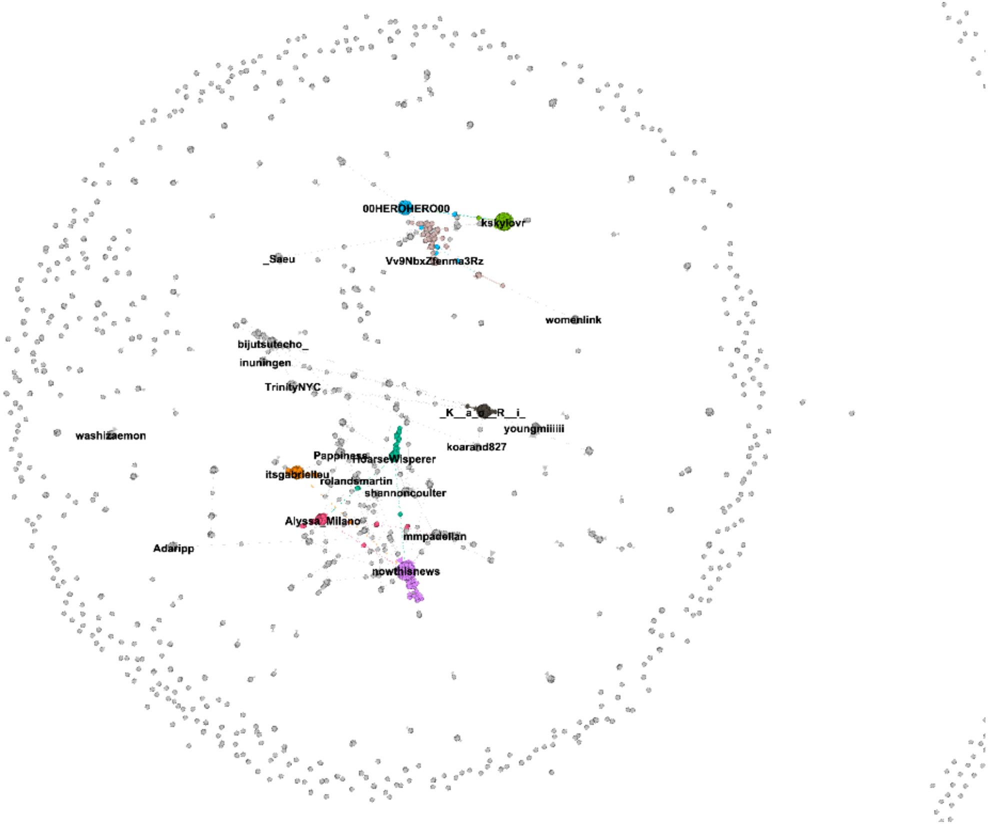


Figure 53: July 6th, 2018 Retweet Network

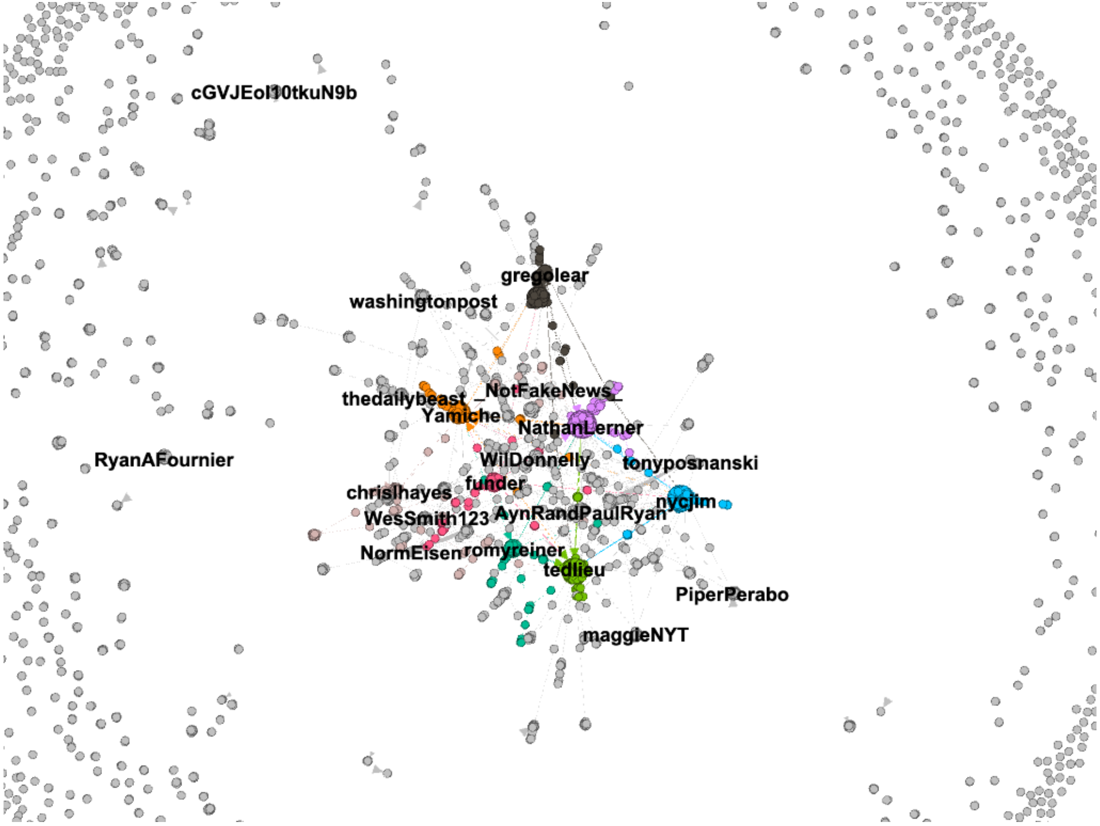


Figure 54: August 20th, 2018 Retweet Network

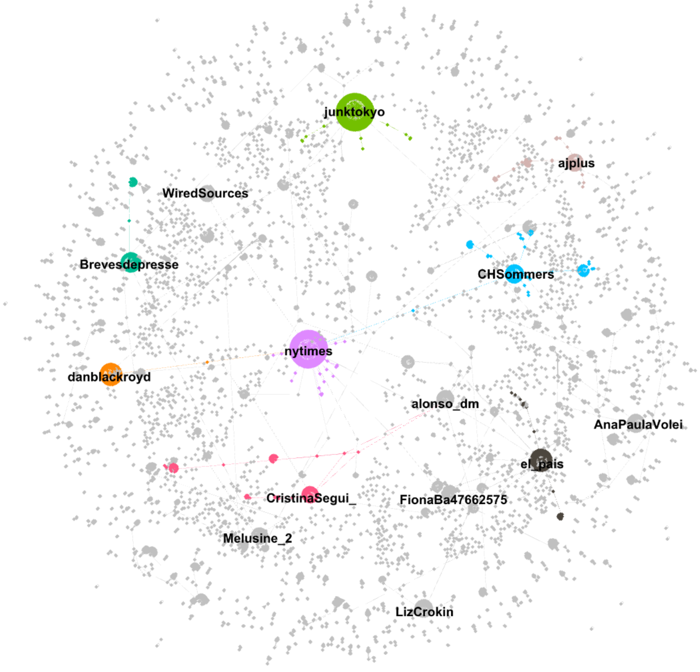


Figure 55: September 9th, 2018 Retweet Network

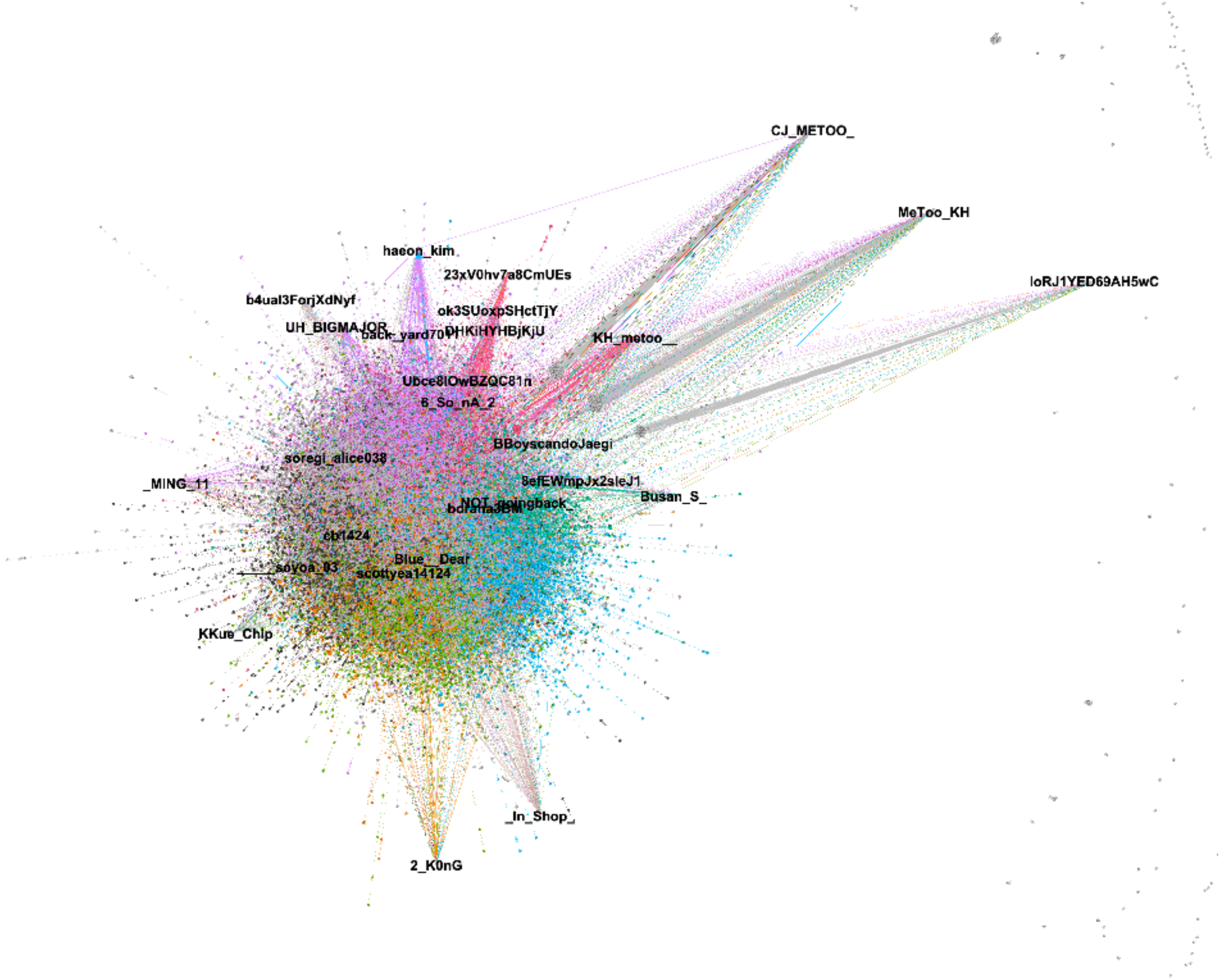


Figure 56: September 16th, 2018 Retweet Network

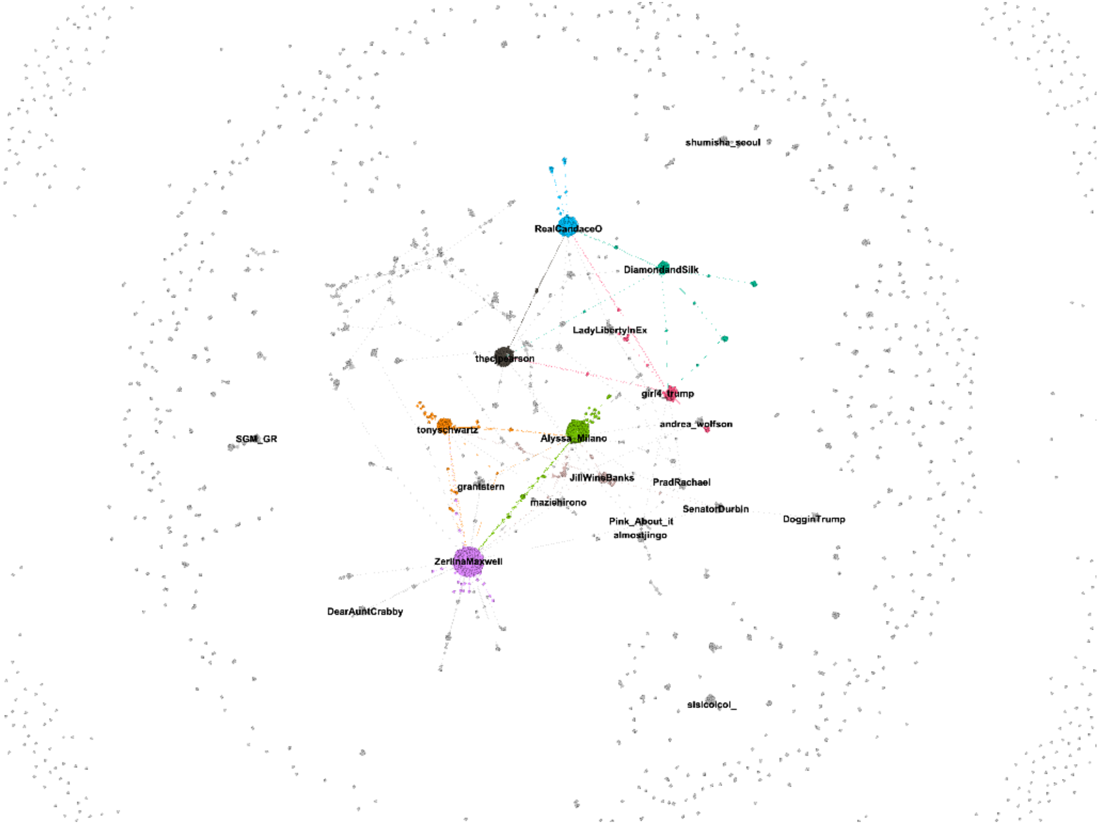


Figure 57: September 27th, 2018 Retweet Network

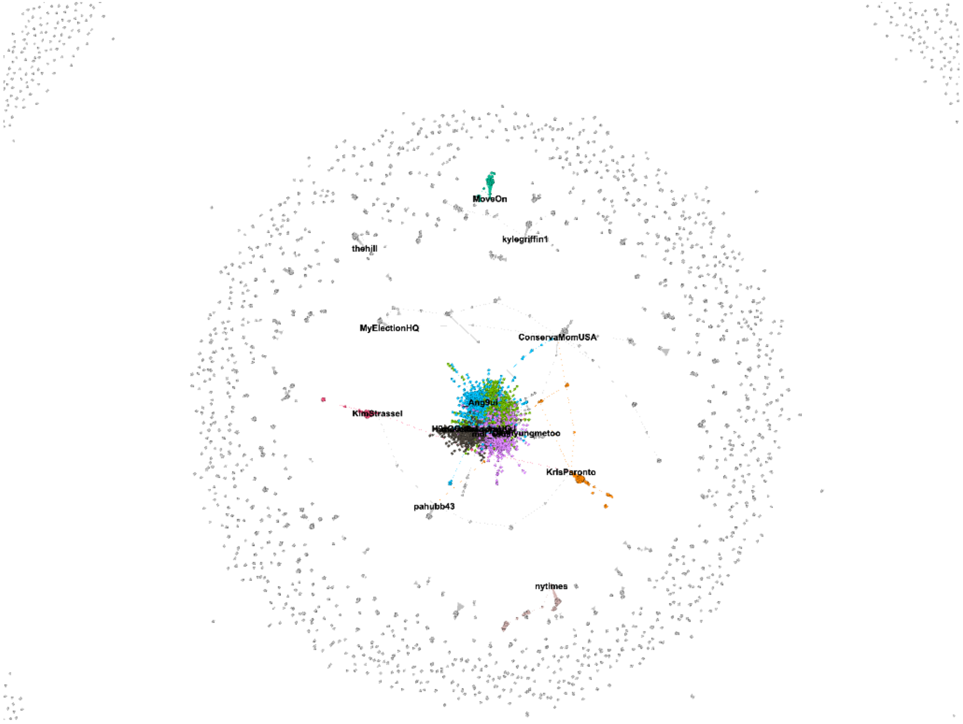


Figure 58: October 6th, 2018 Retweet Network

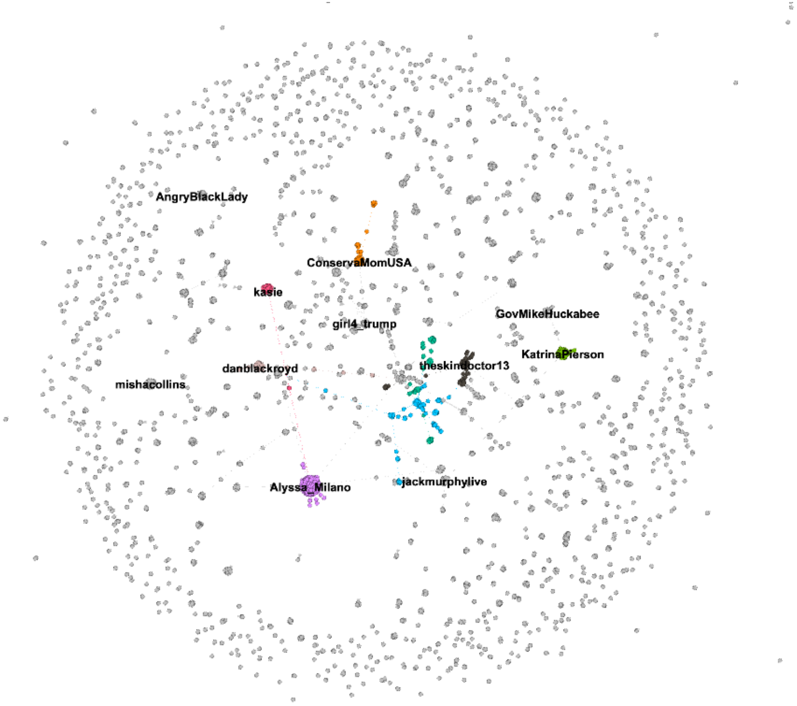
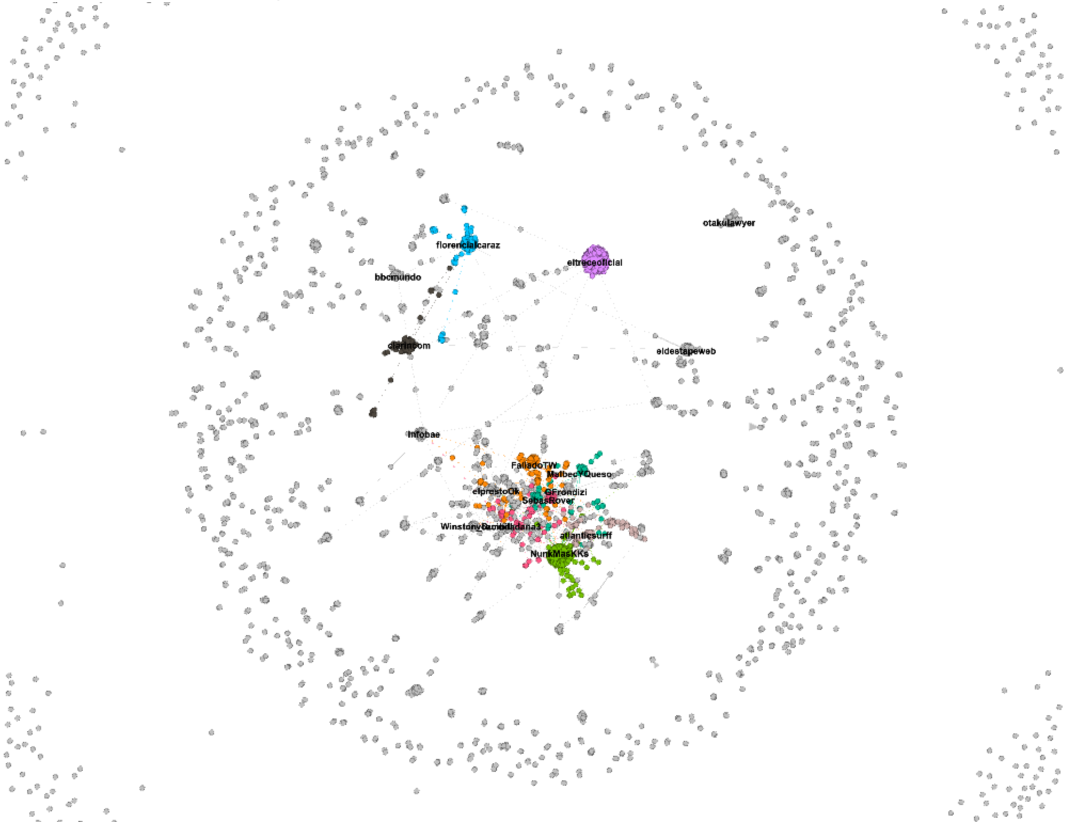
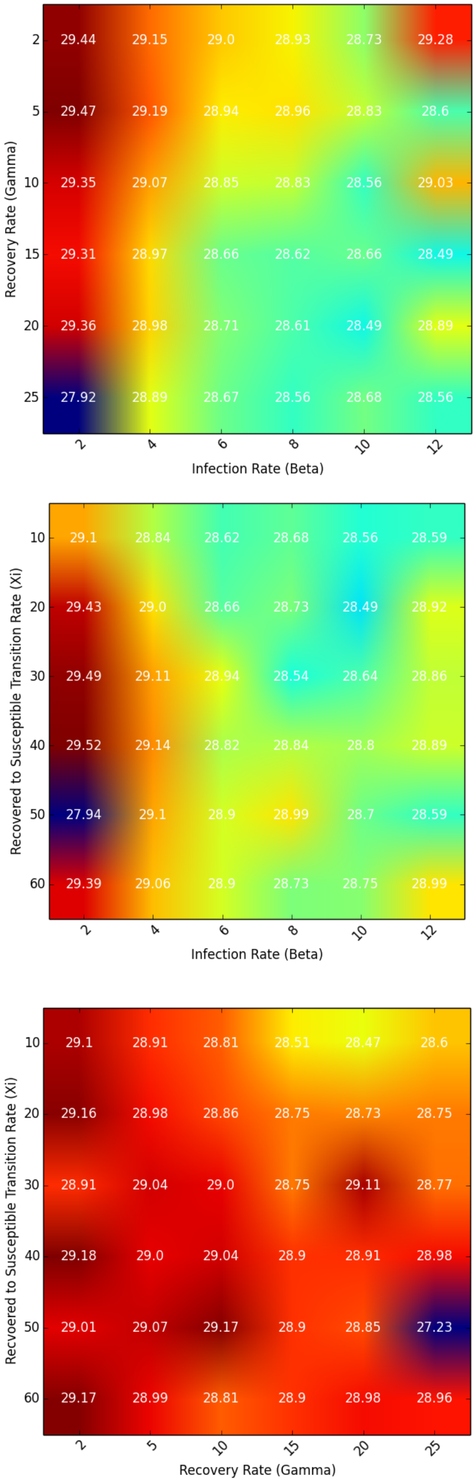


Figure 59: December 11th, 2018 Retweet Network



Language	Twitter Language Code
English (default)	en
Arabic	ar
Bengali	bn
Czech	cs
Danish	da
English (British)	en-gb
German	de
Greek	el
Spanish	es
Persian	fa
Finnish	fi
Filipino	fil
French	fr
Hebrew	he
Hindi	hi
Hungarian	hu
Indonesian	id
Italian	it
Japanese	ja
Korean	ko
Dutch	nl
Norwegian	no
Polish	pl
Portuguese	pt
Romanian	ro
Russian	ru
Swedish	sv
Thai	th
Turkish	tr
Ukrainian	uk
Urdu	ur
Vietnamese	vi
Chinese (Simplified)	zh-cn
Chinese (Traditional)	zh-tw

Figure 60: Heat maps of given sets of parameters for modeling the movement with an SIRS model. The optimal error minimizing parameters are $\beta = 4, \gamma = 25, \xi = 50$, and the minimum mean squared error is 20.5124 with a standard deviation of 13.90875.



All python code used in the data analysis and building of the simulation can be found online on github at:

https://github.com/brookeistvan/seniorthesis.git

References

[1] Google search popularity for me too movement.

[2] Tom Rosenstiel Amy Mitchell. Top 25.

[3] Claire Cain Miller Denise Lu Ash Ngu Jugal K. Patel Audrey Carlsen, Maya Salam and Zach Wichter. metoo brought down 201 powerful men. nearly half of their replacements are women.

[4] Albert-László Barabási. *Network Science*. 2016.

[5] Pablo Barberá, John T Jost, Jonathan Nagler, Joshua A Tucker, and Richard Bonneau. Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological science*, 26(10):1531–1542, 2015.

[6] Pablo Barbera, Ning Wang, Richard Bonneau, John T. Jost, Jonathan Nagler, Joshua Tucker, and Sandra Gonzalez Bailon. The critical periphery in the growth of social protests. *PLoS One*, 10.

[7] Luís M.A. Bettencourt, Ariel Cintrón-Arias, David I. Kaiser, and Carlos Castillo-Chávez. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications*, 364:513 – 536, 2006.

[8] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics*, 2008.

[9] Michael D Conover, Emilio Ferrara, FilippoMenczer, and Alessandro Flammini. The digital evolution of occupy wall street. *PLoS One*.

[10] Twitter Dev. *Supported languages and browsers*. Twitter.

[11] Pedro Domingos and Matt Richardson. Mining the network value of customers. *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 57 – 66, 2001.

[12] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. The rise of social bots. *Communications of the ACM*, 59(7):96–104, 2016.

[13] Institute for Disease Modeling. Generic model. 2018.

[14] Ryan Gallagher, Andrew Reagan, Christopher Danforth, and Peter Sheridan Dodds. Divergent discourse between protests and counter protests: #blacklivesmatter and #alllivesmatter. *PLoS One*, 13.

[15] Jael Goldfine. The new zealand gunman named youtuber pewdiepie, candace owens, donald trump.

[16] Murali Haran. An introduction to models for disease dynamics. 2015.

[17] Herbert W. Hethcote. The mathematics of infectious disease. *SIAM Review*, 42:599–653, 2000.

[18] Janette Lehmann, Bruno Gonçalves, José J. Ramasco, and Ciro Cattuto. Dynamical classes of collective attention in twitter. *Proceedings of the 21st international conference on World Wide Web, WWW'12*, pages 251 – 260, 2012.

[19] Ben Lopman, Kirsten Simmons, Manoj Gambhir, Jan Vinjé, and Umesh Parashar. Epidemiologic implications of asymptomatic reinfection: a mathematical modeling study of norovirus. *American journal of epidemiology*, 179(4):507–512, 2013.

[20] Katerina Eva Matsa and Elisa Shearer. News use across social media platforms 2018.

[21] Graham B McBride and Nigel P French. A new linear analytical sir model for age-dependent susceptibility and occupation-dependent immune status. In *Proceedings of the 6th WSEAS International Conference on Simulation, Modelling and Optimization, Stevens Point, Wisconsin, USA*, pages 222–226, 2006.

[22] Yamir Moreno, Maziar Nekovee, and Amalio F Pacheco. Dynamics of rumor spreading in complex networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 69, 2004.

[23] Azadeh Nematzadeh, Emilio Ferrara, Alessandro Flammini, and Yong-Yeol Ahn. Optimal network modularity for information diffusion. *Physical Review Letters*, 113(8), 2014.

[24] Mary Pflum. A year ago, alyssa milano started a conversation about metoo. these women replied. *NBC News.*

[25] Pablo Piedrahita, Borge-Holthoefer, Javier, Yamir Moreno, and Sandra González-Bailón. The contagion effects of repeated activation in social networks. *Social Networks*, 54:326 – 335, 2018.

[26] NAdia Sayej. Alyssa milano on the metoo movement: 'we're not going to stand for it any more'.

[27] Elizabeth Sebastian and Priyanka Victor. A sir epidemic model with primary immunodeficiency. In *Smart Computing and Informatics*, pages 395–402. Springer, 2018.

[28] Rozina Sini. #metoo, #takeaknee and #covfefe: Hashtags that dominated in 2017. *BBC News.*

[29] Allison Stadd. 10 top accounts to follow on twitter for breaking news. *Adweek.*

[30] Amy Mitchell Tom Rosenstiel. How mainstream media outlets use twitter.

[31] TwitterIR. Q4 and fiscal year 2018 letter to shareholders.

[32] Duncan Watts. A simple model of global cascades on random network. 99(9):57 – 66, 2002.

[33] Jiaocan Wu, Ruping Du, Yingying Zheng, and Dong Liu. Optimal multi-community network modularity for information diffusion. *World Scientific: International Journal of Modern Physics C*, 27(8), 2016.