The Multiple Exposure Mechanism Exploration (MEME) of Social Reinforcement on Reciprocal Twitter Networks

Samantha Fox and Phil Nguyen
CS302 – Final Project
December 11, 2020
The University of Vermont
Burlington, VT USA

Abstract— ***The spread of ideas remains an important phenomenon to understand due to their impact on how they shape the development of beliefs, culture, and society. Progress has been made on several fronts in understanding how they spread using social contagion models, yet the exact mechanisms remain elusive. We extend the traditional SEI (susceptible -> exposed -> infected) disease contagion model by adding multiple exposure compartments (SE_3I) in order to implement and analyze the influence of social reinforcement on idea spread using a random network and a Twitter network of viral memes. Our model produces similar growth behaviors to those of various Twitter memes, but it remains unclear how much of an effect social reinforcement has.***

I. Introduction

There are multitudes of modalities for information to travel and reach people, and somehow certain ideas become adopted around the globe while others remain in a single community. Understanding the mechanisms behind the spread of ideas is particularly important in our current global climate, where scientists and health officials are using all communication modalities possible to convey the severity of the COVID-19 pandemic. If the mechanism behind the spread of ideas could be identified, we could leverage this mechanism to effectively inform individuals around the world. Thus, we are using this platform to introduce a new model that allows us to further explore possible mechanisms of idea spread.

In network science, spreading ideas are often defined as social contagions. The spread of such social contagions is thought to be influenced by social reinforcement. Social reinforcement is defined as the repeated exposure to ideas or behaviors from people in your community that is believed to lead to increased adoption of such ideas or behaviors [1-3]. If social reinforcement requires multiple exposures to infect a susceptible node, then it must be modeled as a complex contagion rather than a simple contagion.

Data from online social media sites such as Facebook, Twitter, or Instagram can be used to analyze the spread of ideas and the effect of social reinforcement. Centola used a manufactured online community and showed that social reinforcement (more than one exposure to a behavior) increases the likelihood that individuals adopt a behavior [1]. Weng et al. used data from twitter to predict viral memes [2]. Weng et al. defined the strength of social reinforcement in a meme as the average number of exposures prior to meme adoption, and found that viral memes did not require social reinforcement while non-viral memes required more exposures to spread and thus concluded that viral memes spread like simple contagions [2]. However, findings from Hébert-Dufresne et al. suggest that all real contagions should be modeled as complex contagions [4]. As real contagions interact with countless other unknown contagions, they must be modeled as interacting contagions, and it was found that interacting simple contagions are indistinguishable from complex contagions despite differing driving mechanisms [4].

To account for other (interacting) factors that may influence someone's decision to adopt or not adopt an idea or behavior, Huo and Chen 2020 used an ISR (ignorant, spreader, stifler) model to consider the influence of internal (scientific knowledge) and external (positive and negative social reinforcement) factors on rumor propagation [3]. It was found that scientific knowledge had an effect on the overall propagation of rumors. Given that finding and the knowledge gleaned from [4], it appears that there are countless internal and external mechanisms that influence the spread of ideas. Unfortunately, we are unable to distinguish one from another if looking at the resulting propagation data as a whole. Thus, we cannot determine the underlying mechanisms of social reinforcement without controlling for all possible and potentially confounding interactions. Although it could be argued that Centola's 2010 experiment accomplished this by creating an isolated social network, that experiment needs to be scaled to see if the findings hold for scale-free networks and to other more integrated (less isolated) networks [1].

It is therefore important to exhaustively search for and develop methods to understand the mechanisms of idea propagation. Centola found that redundant signals significantly increased the likelihood of adoption and that not every redundant signal had the same effect: the second signal was significantly more impactful than the first, the third signal was more impactful than the second, and signals after the third had no more significant impact [1]. Thus, we are investigating the various states of exposure using a modified contagion model, SE_3I (susceptible, exposure 1, exposure 2, exposure 3, infected). We will use our SE_3I model to run several simulations on a configuration model with a power law distribution, and then on a network model containing meme adopters from a reciprocal Twitter follower network collected by [2]. We will run several simulations using the configuration network to understand the effect of multiple exposures and to compare overall propagation trends against real meme propagation data. We will then run the same simulations on networks from various viral memes. We hypothesize that our multi-exposure state model will provide the ability to better match observed meme propagation data whose propagation rate changes overtime.

II. METHODS

To investigate the spread of ideas, we are modeling the various exposure states of a meme spreading through Twitter's social network using a multi-compartment exposure model, SE_3I (Figure 1). In this model, all nodes begin as susceptible (ignorant of the meme) and can follow these steps: (1) susceptible node i can start a new meme spontaneously at rate *a0*, or can become exposed to an existing meme if at least one of its neighbors, i', is infected (meme adopter) with a probability beta that i sees the existing meme from i'; (2) the exposed node i is now in a hesitating state (exposed 1). From exposed 1, node i can either spontaneously adopt the meme by tweeting it at rate *a1*, or get exposed to the meme again from a neighbor, i', and become exposed a second time (exposed 2); (3) the exposed 2 node can now either spontaneously adopt the meme at rate *a2*, or can get exposed to the meme a third time and become exposed 3. (4) From exposed 3, the node spontaneously adopts the meme at rate *a3*.

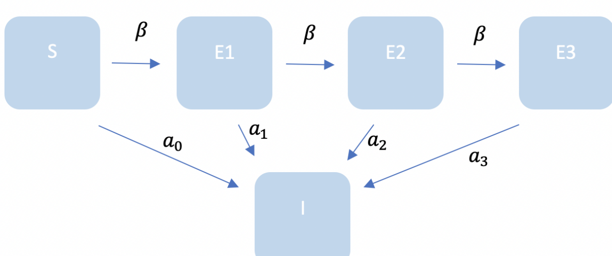


Figure 1: SE_3I (susceptible, exposed 1, exposed 2, exposed 3, infected) compartment model schematic. The compartments are defined as: S is a susceptible state where the node is considered to be ignorant of the meme; E1 is exposed 1 state where the node has been exposed to the meme one time; E2 is exposed 2 where the node has been exposed to the meme twice; E3 is exposed 3 state where the node has been exposed to the meme 3 times; I is the infected state where the node has adopted the meme by tweeting it. The rates are defined as: Beta is the probability that a node sees a meme tweeted by one of its neighbors; a0 - a3 is the rate of transmission from each respective state to I.

To characterize the overall system dynamics from the compartment model, we created the discrete difference equations below:

$$\frac{dS}{dt} = -\beta SI - a_0 S$$

$$\frac{dE_1}{dt} = \beta SI - a_1 E_1 - \beta E_1 I$$

$$\frac{dE_2}{dt} = \beta E_1 I - a_2 E_2 - \beta E_2 I$$

$$\frac{dE_3}{dt} = \beta E_2 I - a_3 E_3$$

$$\frac{dI}{dt} = a_0 S + a_1 E_1 + a_2 E_2 + a_3 E_3$$

To characterize the dynamics of the SE_3I model on a heterogenous network, we derived the mean-field equations below:

$$\dot{S_k} = -\beta k S_k \theta - a_0 S_k$$

$$\dot{E_{1k}} = \beta k S_k \theta - a_1 E_{1k} - \beta k E_{1k} \theta$$

$$\dot{E_{2k}} = \beta k E_{1k} \theta - a_2 E_{2k} - \beta k E_{2k} \theta$$

$$\dot{E_{3k}} = \beta k E_{2k} \theta - a_3 E_{3k}$$

$$\dot{I_k} = a_0 S_k + a_1 E_{1k} + a_2 E_{2k} + a_3 E_{3k}$$

These heterogeneous mean-field equations were implemented in python code (see attached final_project_firstdraft.py), and initial simulations were run using a configuration model with a power law degree distribution to model a social network, and then using network models created from meme adoption data from twitter.

III. RESULTS

<u>Configuration Model</u>: We created a configuration model with a power law degree distribution to run model simulations with our differential equations. We altered the parameters of our model to observe the combined effects of each parameter on meme propagation. By altering the adoption parameters, we can use our model to simulate both social reinforcement, where each exposure is assumed to increase the probability of meme adoption, and the effect of first exposure. The effect of first

exposure is demonstrated when *a0* is small compared to *a1* such that the probability of adopting a meme increases after being aware of the meme. Parameters that do not model social reinforcement or first exposure are parameters that take the form of [*a, a, a, a*], where all adoption parameters are equal. Figure 2 shows examples of all parameters being equal. It was found that each time the simulation was run with parameters in the form of [*a, a, a, a*] (no social reinforcement and no effect of first exposure), the resulting graph showed exponential growth of the *I* term.

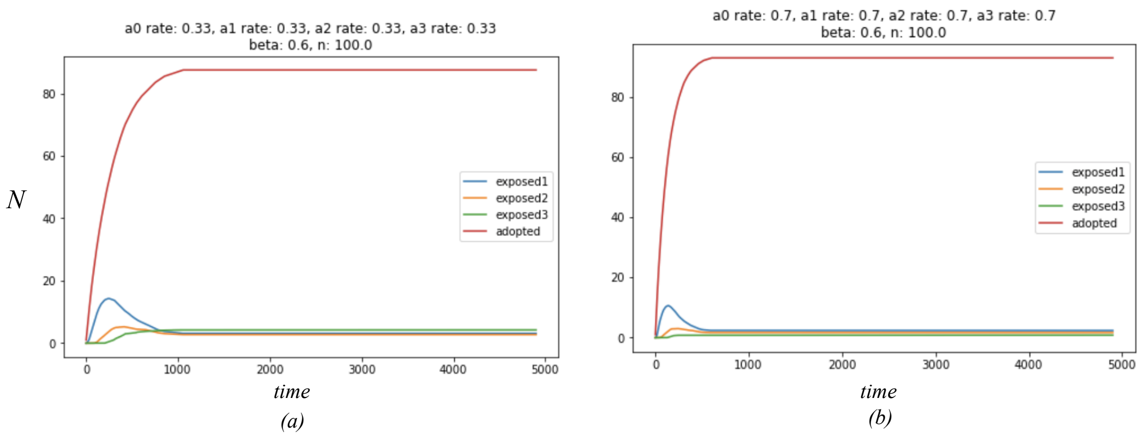


Figure 2: Both a and b are graphs of the number of nodes (N) in each of the exposed states and the adopted state throughout the simulation of the configuration model. These simulations were run with one random adopted seed node. (a) alpha parameters are all equal to 0.33 and beta is 0.6. (b) alpha parameters are all equal to 0.7 and beta is 0.6. T = 50, time-step size = 0.01.

Alpha changes the rate at which nodes become infected. As shown in figure 2, beta remains the same as alpha increases from (a) to (b), also increasing the rate of meme adoption. When all alphas are equal, the greater the alpha, the greater the slope of the growth rate curve.

Figure 3 is an example of adoption parameters $a0 = a1 = a2 = 0$ and $a3 = 0.5$, such that the meme is only adopted at the third exposure. When all *a0-a2* are 0, a high beta is required to observe dynamic behavior. Figure 3 exhibits our model's structure well in that we see an initial delay (flat before exponential growth) caused by the need to move through multiple compartments prior to adoption. We then see exponential growth caused by a jump in adoption probability from 0 to 0.5 at *a3*. In this case, most nodes ended up in the adopted state because of beta's value of 1, and the nodes that did not, remained in exposed 3. Therefore, both alpha and beta contribute to the growth rate and beta contributes to the overall propagation of the meme (i.e. how many people see it). In figure 3, we also observe a few inflection points rather than a smooth curve which is further observed in figure 4.

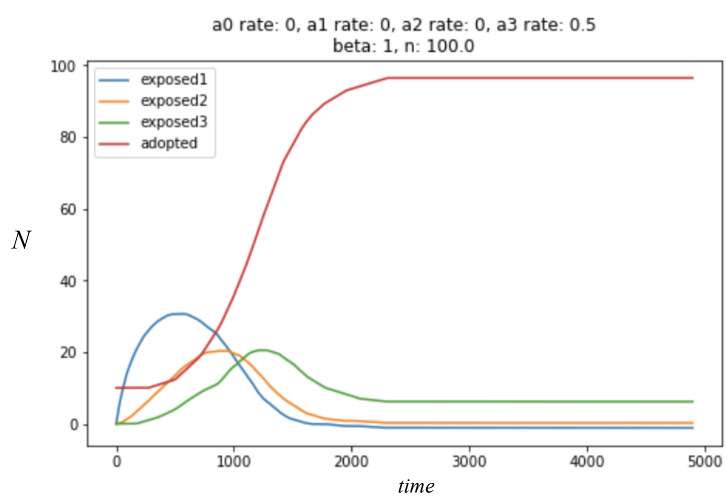


Figure 3: A graph of the number of nodes in each exposed state and the adopted state resulting from the configuration model with 10 random adopted seeds. T = 50, time-step size = 0.01.

To simulate social reinforcement, a variety of parameters that increased from *a1* to *a3* were tested. Figure 4 below shows a simulation using two different sets of parameters. The first simulation in figure 4a uses parameters based on Centola's findings of social reinforcement, while the second in figure 4b uses parameters that do incorporate first exposure but not social reinforcement.

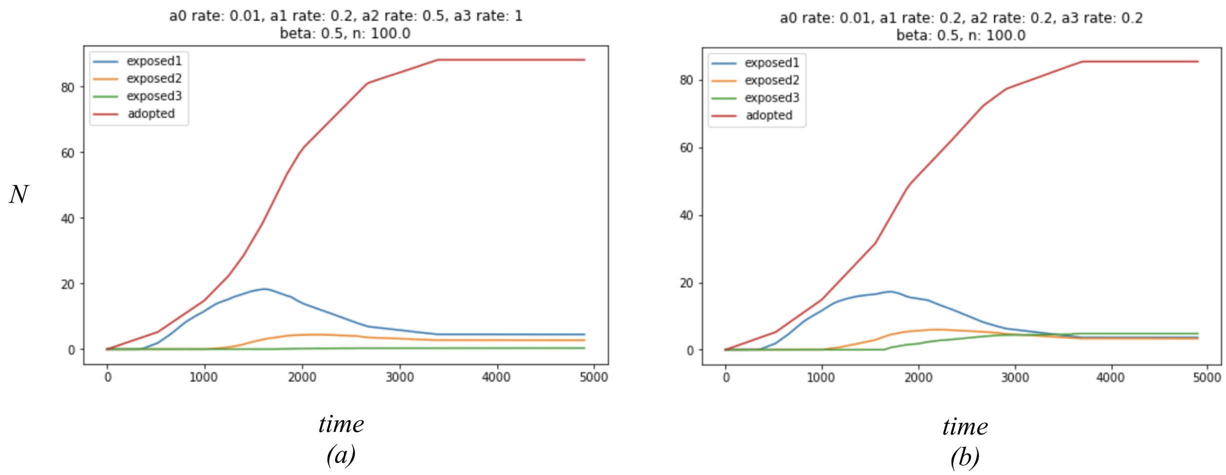


Figure 4: Both figures are simulations from our configuration model. Neither simulation had any seed nodes. (a) increasing alpha parameters (b) a1 - a3 are set to be equal. T = 50, time-step size = 0.01.

We see that the combined effect of the ability to adopt a meme from each state (S, EI, E2, or E3) results in a variety of different growth rates (e.g. linear combined with exponential growth) observed in the overall rate of adoption. We also see that the two simulations produce very similar graphs, which is not what we expected. If social reinforcement is a driving mechanism for idea spread, we would expect the graph without social reinforcement to fail to propagate as far as the graph with social reinforcement. It is possible that each exposure does not necessarily increase the chance of adopting a meme, but rather each exposure gives you another chance to adopt a meme.

<u>Meme Network Model</u>*:* Network structure has a significant effect on the dynamics of behavior diffusion, suggesting that topological features such as degree distribution may also have an effect [2]. We thus ran simulations with networks constructed from meme use data from Twitter. The meme data was sampled from public tweets between March 24, 2012 to April 25, 2012 from [2]. The data contains a reciprocal follower network where nodes are twitter accounts and edges are the connections between two nodes that follow each other. The data contains a sequence list of various hashtags with adopters (i.e. tweeters who use the hashtag) in chronological order of adoption.

We started our meme analysis by looking at memes with different propagation profiles. Figure 5 shows the propagation of each of these memes based on the number of adopters. Wang et al. defined #ThoughtsDuringSchool to be viral and the internet has boasted #kony2012 to be the most viral meme of 2012 (based on unknown criteria). However, these memes follow very different propagation trajectories and potentially have different propagation mechanisms.

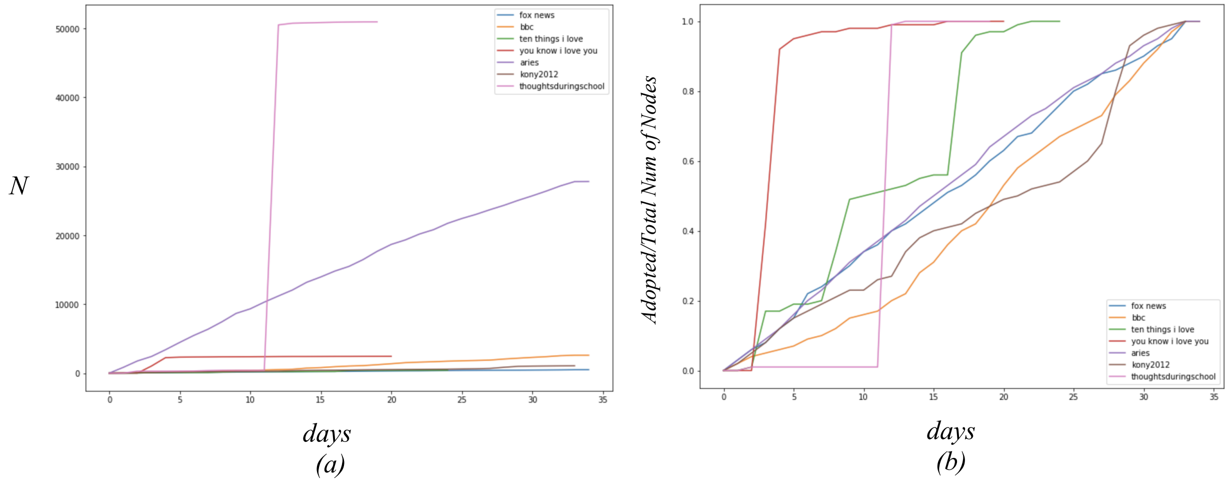
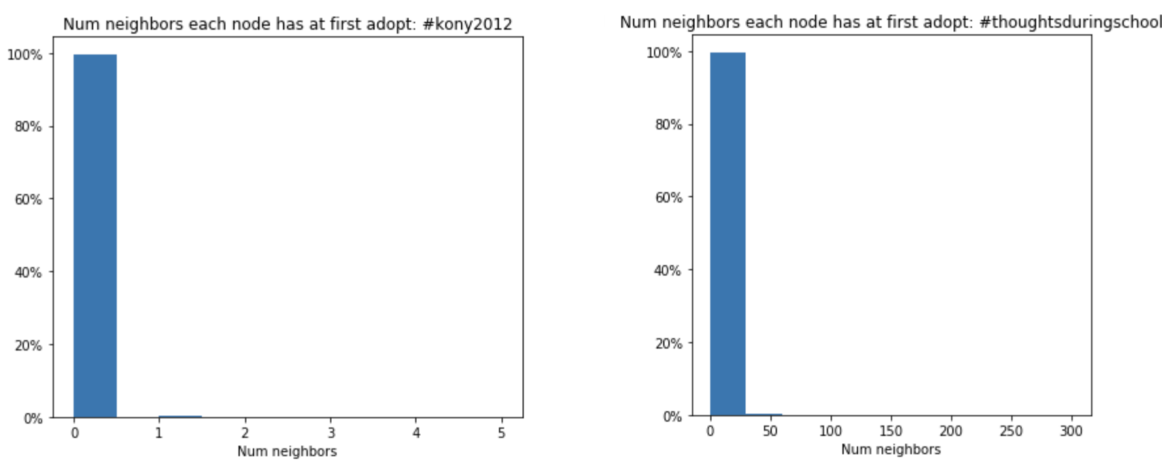


Figure 5: (a) The total number of meme adopters over one month, (b) and the total proportion of those meme adopters over the same time span.

We were interested in investigating if the memes with different adoption rates could be due to the average number of exposures to the meme in that network. We hypothesized that memes like #ThoughtsDuringSchool and #youknowiloveyou would not require social reinforcement because they could be considered less trivial in that they don't make a statement about your personality and the content of the meme is not important, so they could quickly be adopted without hesitation. It was previously found that such viral memes don't require social reinforcement [2], so we expected that most people would adopt the memes without requiring multiple exposures. Conversely, for memes that make a personal statement and might require multiple exposures before adoption such as #kony2012, #bbc, and #foxnews, we expected to see multiple exposures prior to adoption. Yet this is not what we observed.

Figure 6 shows a histogram for each meme with number of exposures on the x-axis, defined as number of infected neighbors at the time of adoption, and percentage of nodes in the network on the y-axis. Figure 6 suggests that if social reinforcement was a mechanism involved in propagating these memes, it may not be present in this reciprocal follower network since a large majority of the nodes did not have neighbors in the adopted state at the time of meme adoption. This may be a consequence of the network structure because it is a reciprocal follower network. It also is possible that social reinforcement could be present in the retweet or mention networks, which are not analyzed here, or via sources outside of Twitter.



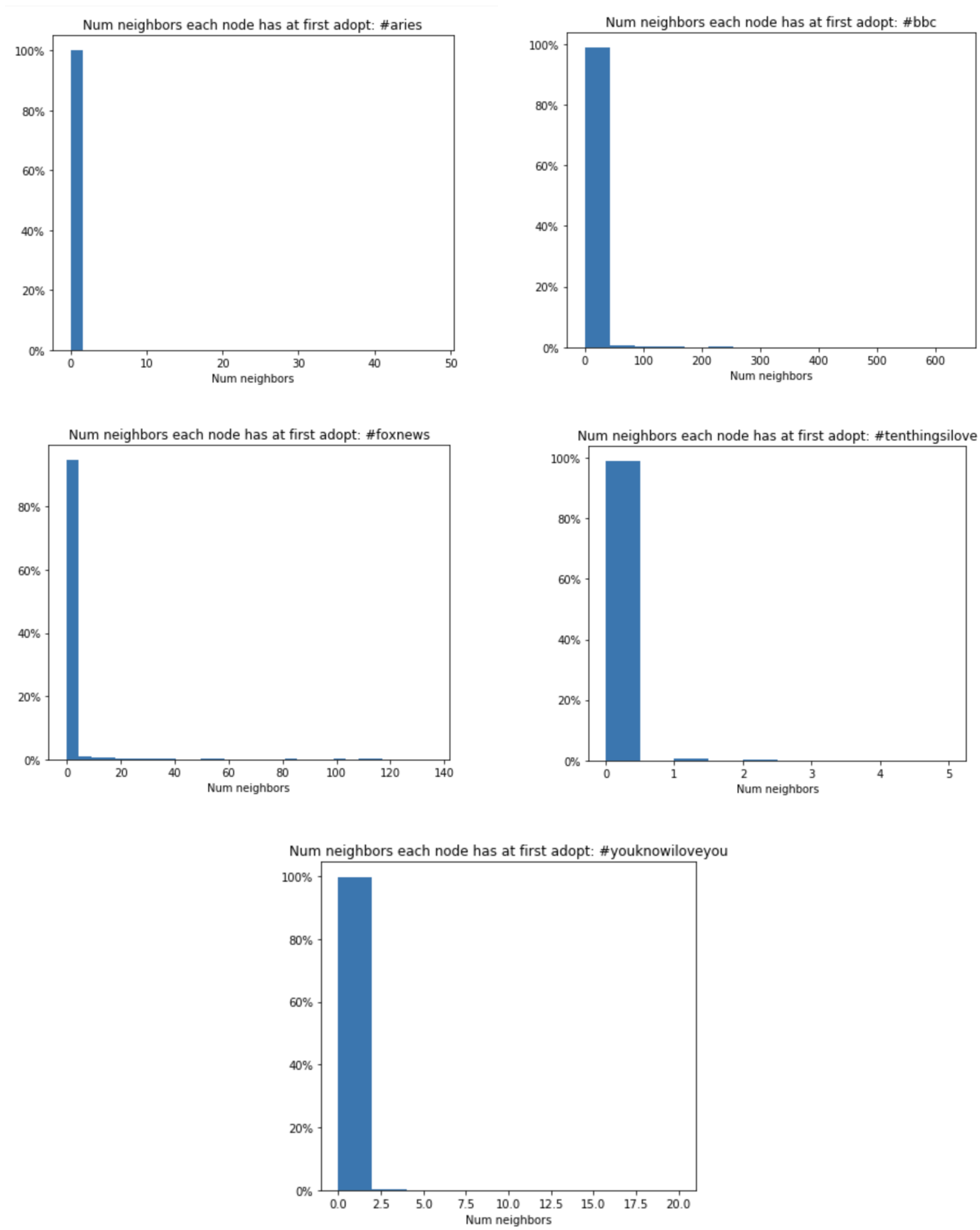
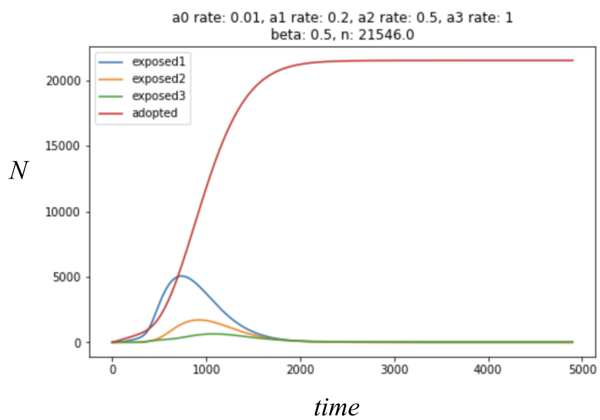


Figure 6: The number of infected neighbors nodes had when they adopted the meme. The number of infected neighbors is on the x-axis (where number of infected neighbors at time of adoption = number of exposures at adoption), and percentage of nodes in the network on the y-axis.

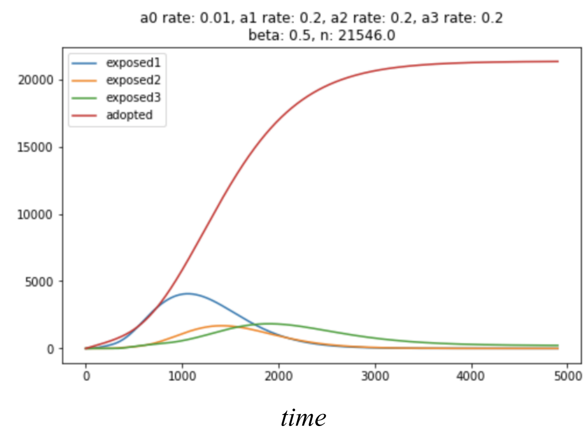
TABLE 1: MEME NETWORK STATISTICS

Meme	Number of Nodes	Number of Edges	Average Degree	Average Clustering Coefficient
bbc	14449	25140	3.479	0.075
foxnews	2492	9024	7.242	0.196
aries	167140	362673	4.339	0.064
thoughtsduringschool	240050	913594	7.611	0.124
tenthingsilove	6961	7202	2.069	0.031
kony2012	21546	22982	2.133	0.014
youknowiloveyou	51495	65488	2.543	0.028
Configuration model	100	97	1.94	0.014

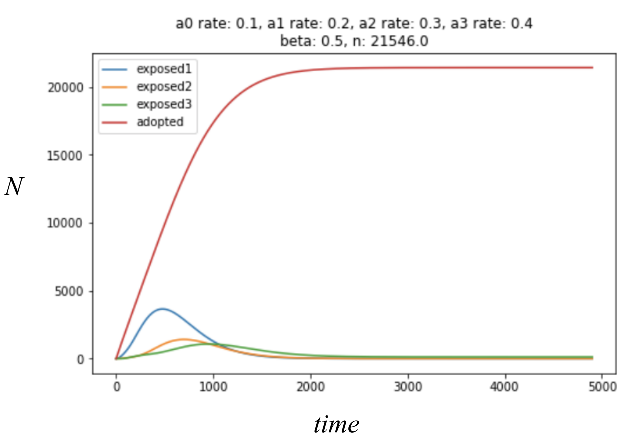
Simulations were then run on all seven constructed meme networks. Propagation results varied slightly between meme networks, but generally tended toward exponential rates of adoption. Simulations took a long time to run due to network size, so given time constraints, comparisons between the meme networks were not possible. In figure 7 below, we present results from simulations on the #kony2012 network because it has the same clustering coefficient as the configuration model.



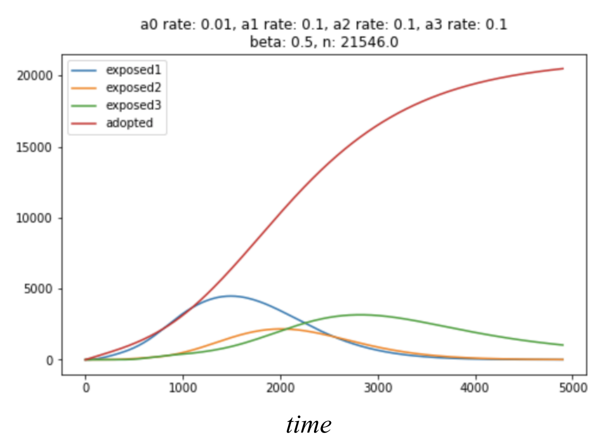
(a)



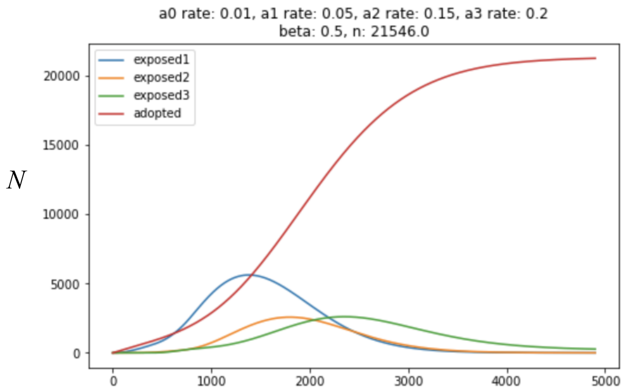
(b)



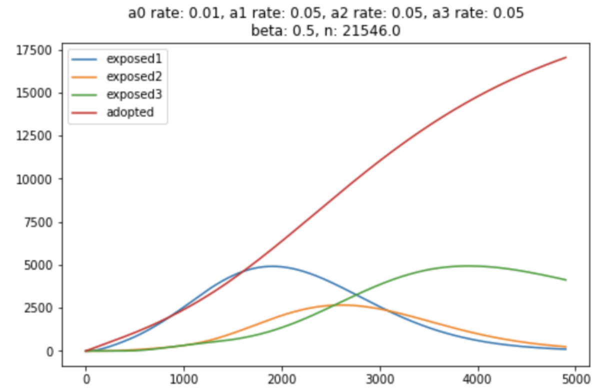
(c)



(d)



time
(e)



time
(f)

Figure 7. Simulations using different alpha parameters on the network constructed from the #kony2012 meme. Social reinforcement is only present in the left column, and not the right. T = 50, time-step size = 0.01.

IV. Discussion

Using the configuration model, it was found that when the effect of exposures on meme adoption was not included in the model parameters, the resulting rate of adoption was exponential every time. Weng et al. found that viral memes such as #ThoughtsDuringSchool propagate without social reinforcement. In figure 5, we see that #ThoughtsDuringSchool grows exponentially, and results from figure 6 suggest that social reinforcement is not coming from the reciprocal follower network, yet we do not know if social reinforcement is coming from another source. While this is not enough data to make any suggestions, it may be worthwhile to further investigate the relationship between meme adoption rate and the effect of exposures.

There are, of course, many cases where social reinforcement in the configuration model also leads to exponential growth such as when beta or alpha are high. However, while it can be realistic for beta to be high (e.g. closer to 1) for someone who is constantly reviewing their twitter feed, it is not realistic for alpha to be high. One must think of alpha as the probability that you see a meme and decide to adopt it—if you had a 20% chance of adopting every meme you saw, even if adoption only occurred at the third exposure, you would be adopting several memes everyday.

While still containing exponential growth periods, using social reinforcement parameters inline with Centola's findings typically provided reduced exponential trends or linear-like trends. One such example is shown in figure 4a. However, while figure 4a models social reinforcement, figure 4b, which has the same *a0* and *a1* parameters as figure 4a, does not. Yet figure 4b exhibits an almost identical adoption rate as 4a. This occurrence begs the question: is social reinforcement increasing the probability of adoption or is it providing an additional opportunity to adopt which increases the cumulative probability of adoption? This question should be further investigated.

Although observations from the meme network simulations did not initially appear to carry the same observations as the configuration model simulations, similarities in network dynamics are present. Results from the #kony2012 network simulations are presented in figure 7 because they were expected to most closely match the configuration model due to equal average clustering coefficient. #kony2012 can be visually compared to the configuration model by comparing figures 7a and 7b (#kony2012 simulations) against figures 4a and 4b (configuration simulations) as they have the same two sets of parameters, respectively. Differences observed include the smoothness of the adoption curves in the #kony2012 model compared to the inflection points observed in the configuration model. Additionally, in terms of proportion of infected individuals, it appears that social reinforcement had little effect on #kony2012 while it had a minor effect on the configuration model. The percentage of infected nodes in 7a was 99% (social reinforcement) compared to 98% in 7b (no social reinforcement), while the percentage of infected nodes in 4a was 92% compared to 84% in 4b. These differences may be due to #kony2012's greater average degree and greater number of nodes (see table 1 for network statistics). Yet despite these differences, we see similarities as well. Similarities are observed in the effect of social reinforcement on the propagation speed. With both the configuration model and the #kony2012 model, we see that the maximum number of adopters is reached slightly faster with social reinforcement than without.

This phenomenon is generally observed in figure 7 where we see that social reinforcement (graphs on the left column of the figure) lead to more exponential growth, reaching the maximum number of adopters quickly, while graphs without social reinforcement (right) tend toward a more linear and slower growth. This is an intuitive outcome given the model structure: when there is a large jump between alpha values, we see an exponential growth in the number of adopted individuals, and when *a1-a3* values are large compared to a0 and beta does not overpower alpha, we are able to observe more linear growth. As this is a structural outcome of our model, we see it in both the configuration model and the meme models.

Another structural outcome of the model that we see in both the meme networks and the configuration model network, is that when the spontaneous adoption from susceptible to infected (*a0*) is very small compared to *a1*, we see delays in the meme adoption rate (figure 7 a, b, d, e, f, g). This is a finding that is worthwhile to further investigate because it is observed in the propagation of memes shown in figure 5. If the difference between a0 and a1 was large, we would observe immediate exponential growth.

There were many limitations to our model investigations presented here. One limitation was that beta likely changes overtime as a meme propagates. In future studies, we suggest that beta changes as the rate of adoption changes. Providing an adaptive beta parameter may better model realistic meme propagation as observed in figure 5. Another limitation is that the meme networks were reduced to only include nodes that adopted the meme. This network reduction was performed to enable us to run simulations under time constraints as the size of the original networks was time prohibitive. However, because network structure has an effect on diffusion dynamics, results on the original networks could be different than those presented here. Additionally, although the Twitter data from [2] included three networks (a reciprocal follower network, reciprocal retweet network, and a reciprocal mention network), we only used the reciprocal follower network, which limits potential interactions that we did not observe through retweets and mentions. Finally, our model is limited to three exposure states. This was designed as such because it has been reported that exposures beyond the third do not increase the probability of adoption [2,5], but it may have limited our analysis.

V. Conclusion

Our model offers a unique perspective to understanding how social ideas spread by modeling separate compartments for each exposure, up to three exposures. With three exposure compartments, we could use our model to simulate the effect of first exposure, and compare the effects of social reinforcement versus no social reinforcement. In doing so, our model produced various types of growth rates, combining linear and exponential growth, which is observed in meme propagation shown in figure 5. Although our model could produce various growth rates, it could do so with a variety of parameters and we are unable to distinguish between a curve produced with social reinforcement and one that was not. For instance, a meme propagation curve that exhibits exponential growth could be due to social reinforcement, a high beta, or all alpha parameters being equal. Therefore, propagation mechanisms are difficult to distinguish. That being said, results from our model provided us the opportunity to question if social reinforcement was driven by an increase in probability due to more exposures as previously thought, or if more exposure simply provides more opportunities to adopt while the probability of adoption remains unchanged. We recommend that future work attempts to distinguish these two potential mechanisms.

References

[1] D. Centola, "The Spread of Behavior in an Online Social Network Experiment," *Science*, vol. 329, no. 5996, pp. 1194–1197, Sep. 2010, doi: 10.1126/science.1185231.

[2] L. Weng, F. Menczer, and Y.-Y. Ahn, "Virality Prediction and Community Structure in Social Networks," *Sci. Rep.*, vol. 3, no. 1, p. 2522, Dec. 2013, doi: 10.1038/srep02522.

[3] L. Huo and S. Chen, "Rumor propagation model with consideration of scientific knowledge level and social reinforcement in heterogeneous network," *Phys. Stat. Mech. Its Appl.*, vol. 559, p. 125063, Dec. 2020, doi: 10.1016/j.physa.2020.125063.

[4] L. Hébert-Dufresne, S. V. Scarpino, and J.-G. Young, "Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement," *Nat. Phys.*, vol. 16, no. 4, pp. 426–431, Apr. 2020, doi: 10.1038/s41567-020-0791-2.

[5] M. Zheng, L. Lü, and M. Zhao, "Spreading in online social networks: The role of social reinforcement," *Phys. Rev. E*, vol. 88, no. 1, p. 012818, Jul. 2013, doi: 10.1103/PhysRevE.88.012818.