Simulating a Fake News Epidemic

Laura Cebollero Ruiz
Carolina Centeio Jorge
Kymry Burwell

Abstract—

Social media has become a large part of everyone's life. It is a strong source of information because of its easy access, fast dissemination and low cost. However, it also enables the wide spread of fake news. Fake news are news with convenient misinformation. The propagation of fake news in social networks is a subject of recent research. What's more, we can see fake news as an epidemic in complex networks. As such, we can try to apply epidemic models to a network, in order to predict the fake news propagation and, consequently, mitigate its impact "immunizing" convenient nodes (users). In this paper we discuss three potential SI models. Our hope is that such models, or future variations of them, can be used as test immunization strategies that can later be used on real social networks in an attempt to slow or prevent the spread of fake news.

Keywords—epidemic, fake news, simulation, SI, SIR, immunity, spread, propagation, network

1 Introduction

Nowadays people constantly consume information from online sources [Shu et al., 2017b]. More than 62% of U.S. adults turn to social media for news; 18% do it often[1]. Social media tends to be more present in our lives than traditional organizations [Shu et al., 2017a]. This happens for two main reasons: it is often less expensive and it is easier to further share, comment and discuss.

Although there are clear advantages on making use of social media, the quality of its content is worse than that of the traditional sources [Shu et al., 2017a]. Social media easily spreads fake news (hoaxes, rumors, conspiracy theories, fabricated reports, click-bait headlines, and even satire [Shao et al., 2017]). Fake news has content that is intentionally false and misleading [Shu et al., 2017b]. These news can be seriously dangerous and harmful for societies [Farajtabar et al., 2017]. In particular, people may be misled by fake news and accept false beliefs. It can also affect how people react to real news. Falsehood diffuses significantly farther, faster, deeper, and more broadly than the truth in many information; the effects are more pronounced for false political news than for false news about terrorism, natural disasters, science, urban legends, or financial information [Vosoughi et al., 2018]. Furthermore, it has been shown that once absorbing misinformation from fake news, individuals are less likely to change their beliefs even when the fake news are debunked. If the idea in the original fake news is especially similar to individuals' viewpoints, it will be even harder to change their minds [Ecker et al., 2010]. For these reasons, digital misinformation has become so pervasive in online social media that it has been listed by the WEF (Word Economic Forum) as one of the main threats to human society [Del Vicario et al., 2016]. As such, it is of utmost importance to develop strategies to reduce the impact of fake news.

There are big companies behind social media already trying to mitigate this impact. Facebook's strategy, for example, is providing easy options to users to report[2] a hoax they see in the social

1. News Use Across Social Media Platforms 2016 http://pewrsr.ch/249haJM

2. https://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/

network. They rely on the community to mitigate fake news. After a post being reported as fake news, the post gets a flag as it might contain misinformation to inform the next users that see it. Such direct action on the offending news not only requires a high degree of human oversight, which can be costly and slow, but it may also violate civil rights. A clear disadvantage of the report-and-flag mechanism is the possibility of abuse by adversaries who conveniently report real news as fake news. Thus, a possible alternative is optimizing the performance of real news propagation over the network, ensuring that people who are exposed to fake news are also exposed to real news, so that they are less likely to be convinced by fake news [Farajtabar et al., 2017]. Early detection of fake news is very desirable to restrict the dissemination scope of fake news and prevent the future propagation on social media. Early fake news detection aims to give early alert of fake news, by only considering limited social context within a specific range of time delay of original news posted [Shu et al., 2017b]. This propagation follows some patterns that can be used for fake news detection in social media [Shu et al., 2017a]. On the other hand, fake news is an epidemic which requires a grounded assessment of the mechanism by which misinformation spreads online [Wang, 2017].

Since the spread of fake news can be seen as an epidemic, we can try to apply epidemic models over the networks that have the fake news spreading on. These networks are social networks in which the nodes represent the users and the edges represent the interactions (or connections) between users. There are three well-known important epidemic propagation models, namely SI, SIR and SIS [Brede, 2012]. The SI model (fully mixing susceptible-infected), considers two states of an individual: the susceptible and the infected. Contact with infected individuals causes a susceptible person to become infected and once infected a person cannot recover. There is a probability of an infected infecting a susceptible neighbor per unit time (an approximation) [Pastor-Satorras and Vespignani, 2001]. If we want to consider the node can recover after being infected, we can use the SIR model. In this model we can consider three probabilities: probabilities that vertex i is susceptible, infected, or recovered. After becoming recovered, the nodes cannot change to another state (susceptible of infected). Finally, we can also consider the fact that, after recovering, the node can be infected again; so a node can be susceptible, get infected and return to susceptible state (SIS model). In this method, when a node recovers it becomes susceptible again; there's an infection rate and a recovering rate.

In this paper we present three approaches for modeling the propagation of fake news. With the use of complex network concepts, we implemented an SI and an SIR model for a fake news epidemic. We will be using the dataset that is described in [Shu et al., 2018b]. More specifically, we will be using the *Politifact* dataset and the files we have picked contain information about: (1) A User-User network that relates users that are connected and (2) A News-User network that relates which news have been shared by which users. The main goal is to find a model, along with a steady β parameter, that accurately simulates the spread of fake news over a network. The thought is that if an accurate model can be found, then on future work mitigation strategies can first be tested on the model and then implemented on real social networks in an attempt to reduce the spread of fake news.

The rest of the paper is organized as follows: We summarize the important related work in section 2, describe the dataset in section 3, present our experiments in section 4 to later compare them in 5. Finally, we reach some conclusions in section 6, along with some suggestions for future directions.

2 Related Work

The authors that gathered, provided and continuously update the dataset we are using for this project wrote three papers, them being [Shu et al., 2017a], [Shu et al., 2017b] and [Shu et al., 2018b]. The authors refer how late studies have tried to develop tools for fake news detection and, on the other hand, how hard and complex it is to develop these tools. The main reason for this, the authors state, is that fake news are written with a special intention of misleading the readers, fooling them, which makes them carefully written and easily credible. Since they do not necessarily comply in terms of topic or linguistic metrics, hand-crafted and data-specific textual features are generally not sufficient. Also, they are sometimes anonymous and noisy. Moreover, fake news is usually related to newly emerging events, which may not have been properly verified by existing knowledge bases due to the lack of corroborating evidence or claims. In [Shu et al., 2017a], the authors focus in two main aspects of fake news detection problem: *characterization* and *detection*. They discuss the definitions of fake news in the literature, give an overview of existing fake news detection methods and discuss open issues on the topic of fake news detection in social media. As a relevant definition for our work, the authors define Social News Engagements as a set of tuples $\varepsilon = \{\epsilon_u\}$ to represent the process of how news spread over time, among n users $U = \{u_1, u_2, ..., u_n\}$ and their corresponding posts $P = \{p_1, p_2, ..., p_n\}$ on social media regarding news article a. Each tuple (engagement) will then be $\epsilon_u = \{u_x, p_y, t\}$, meaning that u_x spread article a in post p_y at time t. If $t == NULL$ it is representing the publisher. They refer that propagation-based approaches for fake news detection reason about the interrelations of relevant social media posts to predict news credibility. In fact, [Gupta et al., 2012] explore the possibility of detecting credible events from Twitter feeds using credibility analysis. They used a new credibility analysis model for computing credibility of linked set of multi-typed entities. Apart from exploiting tweet feed content-based classifier results, they develop a PageRank-like credibility propagation using a simple network of tweets, users and events (BasicCA). The Basic Credibility Analyzer (BasicCA) initializes the credibility of different nodes in a network of tweets, users and events using the results from their classifier; Using PageRank-like iterations, this initial credibility is propagated all across the network. Nodes share their credibility with their neighbors only, at each iteration. The authors conclude that credible entity links with a higher weight to more credible entities than to non-credible ones. Then, they enhance BasicCA by performing event graph-based optimization to assign similar scores to similar events. In [Abbasi and Liu, 2013], authors propose to address user credibility to tackle the information credibility problem in social media; they propose the *CredRank* algorithm, which measures the credibility of social media users based on their online behavior. The authors suggest that less credible users are more likely to coordinate with each other and form big clusters, while more credible users are likely to from small clusters. This approach prevents the spread of misinformation generated by less credible users, which attempts to increase the quality of information in social media. Also, [Jin et al., 2014] propose a hierarchical propagation model, detecting sub-events within a news event to describe its detailed aspects; then, for a news event, a three-layer credibility network consisting of event, sub-events and messages can represent it from different scale and reveal vital information for credibility evaluation. Authors also utilize a graph optimization framework to infer event credibilities.

Regarding the mitigation part, [Farajtabar et al., 2017] present the first formulation of fake news mitigation as the problem of optimal point process intervention in a network; also, the first to conduct a real-time point process intervention experiment. By defining a state space for the network, formulating actions as exogenous intensity, and defining reward functions, the authors map the fake news mitigation problem to an optimal policy problem in a Markov decision process (MDP), which is solved by model-based least-squares temporal difference learning (LSTD) specific to the context of point processes. The authors define Event Exposure as a quantitative measure of campaing influence and represent it as a counting process that increases everytime the user or a neighbor performs an activity. Cascades of mutual excitations to occur among many nodes are possible, so non-adjacent users can also contribute to one another's exposure counts, if there is a directed path between them. In the paper, Intervention is the incentive of a subset of users at a certain t time to trigger real news events.

[Vo and Lee, 2018] make use of the fact that there are some people that individually try to mitigate fake news spread through the sharing of information that debunks these fake news.

The authors identify these Twitter users, the *guardians*, and develop a recommender system that will present to the guardians facts they would be interesting on sharing and that would debunk a fake news. In a certain way, we can also say they use credibility to mitigate the problem.

Back to the use of complex networks, [Shu et al., 2018a] define their mitigation strategy as the aim to proactively block target users or start a mitigating campaign at an early stage. They show how network diffusion models can be applied to trace the provenance nodes and paths of fake news. Moreover, they state that the impact of fake news can be assessed and mitigates through network estimation and network influence minimization strategies. The network properties stated as having potential role for the study of fake news are echo chambers, individual users, filter bubbles and malicious users. Echo chamber effect results from social polarization, since people follow like-minded people and thus receive news that promotes their preferred, existing narratives; people are more likely to believe in a source if others have previously believed (social credibility) and users favor information they hear frequently (frequency heuristic). Echo chambers may create segmented and polarized communities. Individual users may play different roles throughout the spread of a fake news, such as persuaders (the ones intentionally spreading fake news), gullible (users that easily believe fake news) and clarifiers (the skepticals). These classifiers can also spread true news which can immunize users against changing beliefs before they are affected by fake news and further propagate and spread true news to other users. Filter Bubbles personalize information to which a user has easy access to; this can create intellectual isolation, since the user may not be exposed to contradicting ideas and amplifies the challenges to dispel fake news. Finally, Malicious Accounts can amplify the spread of fake news, them being social bots, trolls, cyborg users. They consider two main types of networks: Homogeneous and Heterogeneous. Homogeneous networks have the same node and link types, namely, social relations (where all the nodes and edges are of same type, of course), information propagation, etc. On the other hand, heterogeneous networks have a different set of node and like types. When speaking of networks, the authors also refer provenance paths of information, which remain an open problem; they suggest a solution based on the possible paths of information propagation using provenance seeking nodes (the provenance paths). Regarding the identification of persuaders, the authors state that the problem of finding K-leaders is tantamount to the maximization of a nonnegative, nondecreasing, submodular function with cardinality constraint that can be solved by a hill-climbing algorithm with a provable constant approximation. The impact of fake news on social media can be estimated as the number of users who are potentially affected by the news piece, an amount we want to assess and then minimize. Their goal is to stop the information cascade mentioned in [Castillo et al., 2014], limiting the influence of fake news. The problem can be seen as a minimization problem: minimize the number of final infected users by blocking k users. Also, the authors aim to maximize the spread of true news.

As said before, limiting the spread of fake news can be seen as analogous to inoculation in the face of an epidemic. Models of epidemics generally assume that a global parameter describes the probability that a user is infected by a neighbor. This assumption is, however, violated in real-world situations of information exchange where users have varying degrees of willingness to accept information from their neighbors. Thus, Independent Cascade Model (ICM) assumes that each edge has its specific activation probability [Saito et al., 2008]. ICM is denoted as a sender-centric model. At each time t, there's a probability a node u infects a neighbor u'.

Furthermore, the fake news diffusion process has different stages in terms of people's attention and reactions over time, resulting in a unique life cycle different from that of in-depth news [Castillo et al., 2014]. Authors show that it is possible to model accurately the overall traffic articles will ultimately receive by observing the first ten to twenty minutes of social media reactions. The distribution of Tweents or Facebook shares of the News and In-Depth content of Al Jazeera[3] can be seen in Fig. 1, They conclude that News articles describing breaking news events tend to attract

3. www.aljazeera.com

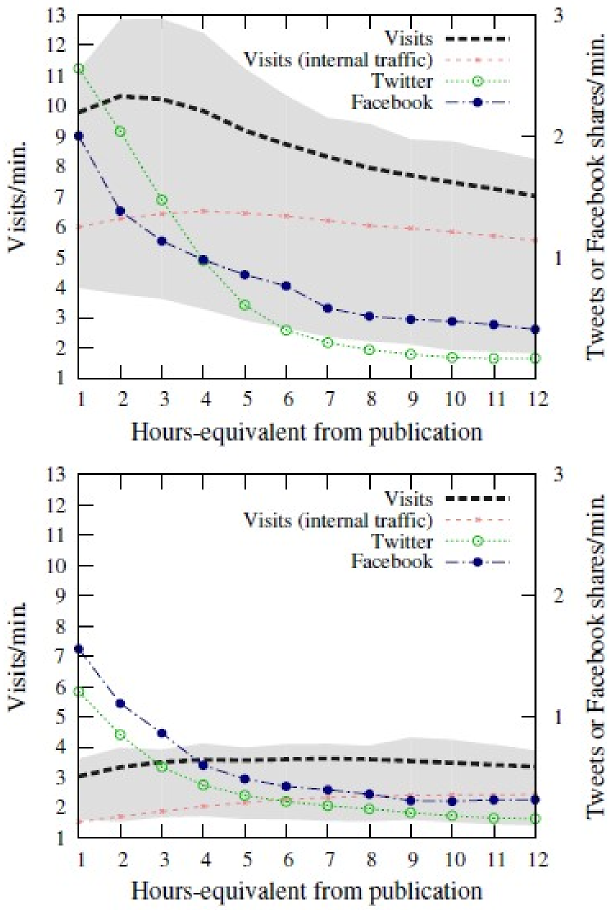

Figure 1: "Visits per minute (left y-axis) as well as Tweets and Facebook shares per minute (right y-axis) for the first 12 hours. For visits, the shaded area covers 50% of the data (quantiles 0.25 to 0.75). Top: average for a News item. Bottom: average for an In-Depth item." from [Castillo et al., 2014]

attention shortly after they are published and so have a shorter shelf-life; they have more repetitive social media reactions, as most users simply repeat the news headlines without commenting on them. However, In-Depth items portraying or analyzing a topic tend to exhibit a longer shelf-life.

[Vosoughi et al., 2018] investigated the differential diffusion of all of the verified true and false news stories distributed on Twitter from 2006 and 2017. The dataset gathers around 126,000 stories tweeted by approximately 3 million people more than 4.5 million times. The authors classified news as true or false using information from six independent fact-checking organizations that exhibited 95 to 98% agreement on the classifications. The authors found that false news was more novel than true news, which suggests that people were more likely to share novel information. Whereas false stories inspired fear, disgust, and surprise in replies, true stories inspired anticipation, sadness, joy, and trust. Contrary to conventional wisdom, robots accelerated the spread of true and false news at the same rate, implying that false news spreads more than the truth because humans, not robots, are more likely to spread it. A very interesting contribution of this work was the study of the evolution of rumor cascades. Part of it can be seen in Fig. 2. Contrarily, [Wang, 2017] provide quantitative empirical evidence of the key role played by social bots in the viral spread of fake news online. Relatively few accounts are responsible for a large share of the traffic that carries misinformation.

[Del Vicario et al., 2016] conclude that for science and conspiracy news a cascade's lifetime has a probability peak in the first 2h (visible in Fig. 3), followed by a rapid decrease. These results suggest that news assimilation differs according to the categories. Science news is usually assimilated; it reaches a higher level of diffusion, quickly, and a longer lifetime does not correspond to a higher level of interest. Although the consumption patterns are similar, cascade lifetime as a function of the size differs greatly (can be seen in Fig. 4); conspiracy rumors are assimilated more slowly and show a positive relation between lifetime and size.

Finally, [Bessi et al., 2015] obtained results that show that polarized communities emerge around distinct types of contents and usual consumers of conspiracy news result to be more focused and self-contained on their specific contents.

3 Dataset

The approach to create a simulation system SI and discover which β fits best the spread of fake news, we have used the dataset *Politifact* from the *FakeNewsNet* project ([Shu et al., 2017b], [Shu et al., 2017a] and [Shu et al., 2018c]).

This dataset contains two main files that we have focused on working on:

- `PolitiFactUserUser.txt`: contains the definition of the User-User network. List

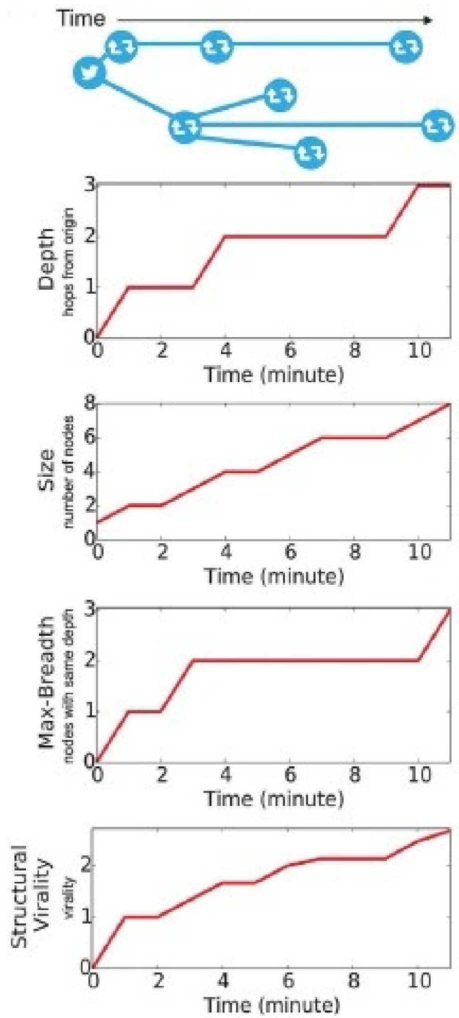

Figure 2: "An example rumor cascade collected by our method as well as its depth, size,maximum breadth, and structural virality over time. “Nodes” are users.” from [Vosoughi et al., 2018]

of edges. This is the network we are going to work on.

- `PolitiFactNewsUser.txt`: contains the definition of the News-User network, relating for each fake news available, which user has shared it.

3.1 Constraints

Using the aforementioned dataset in order to accomplish our goal, we can point out a few problems:

1) We do not know which users have consumed the news. Even if a user has not shared the news URL, they can have consumed it or shared it via another platform. This is not registered and thus, we cannot know the real damage done by the fake news spread.
2) We do not have time steps to know when the fake news has been shared.
3) On the same line, we cannot know which user has infected whom.
4) Consequently, we also cannot know which users first shared the news.

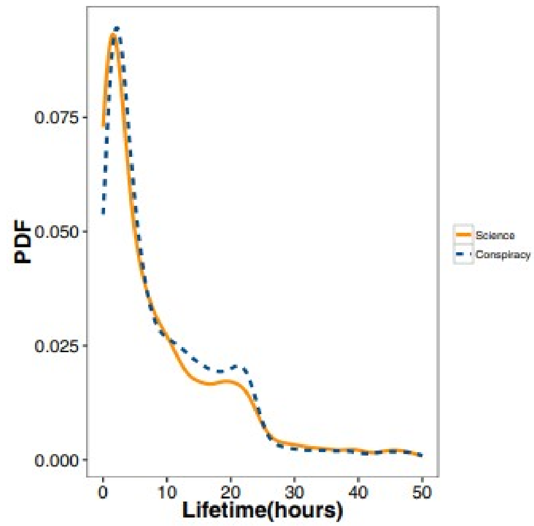

Figure 3: "PDF [probability density function] of lifetime computed on science news and conspiracy theories, where the lifetime is here computed as the temporal distance (in hours) between the first and last share of a post. Both categories show a similar behavior” from [Del Vicario et al., 2016]

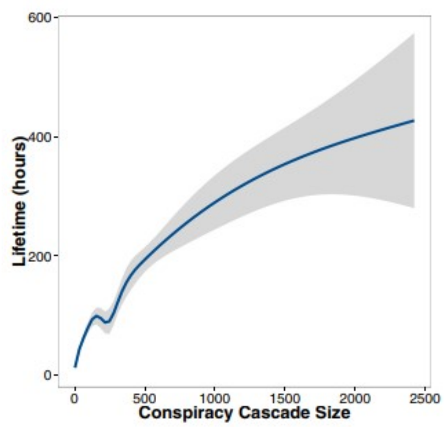

Figure 4: "Lifetime as a function of the cascade size for conspiracy news.” from [Del Vicario et al., 2016]

5) Finally, we do not know whether the data has been collected very far away from the real time event (i.e. 1 year apart from when the news was shared for each user).

3.2 Assumptions

Because of the constraints originated from using said dataset, and for the sake of simplicity, we have assumed the following:

1) A user is infected if, and only if, s/he has been gullible enough to share the fake news among their connections. In other words, having shared the fake news means the user is infected.
2) If a user has not shared the news, he will be deemed as a susceptible individual that may, eventually, share the fake news.
3) We have set, in all models, a maximum time step on which the fake news does not spread anymore. This has been made under the assumption that the fake news dataset has been collected, at least, 2 days after its first appearance on the network. Thus, we are also assuming that after 2 days tops, the fake news spread will not grow anymore.

3.3 Network analysis

Before proceeding onto the Simulation approaches, we first should explore the User-User network. We are going to create a graph G such that each user-user network is bidirectional. Thus, our graph will be undirected.

Said graph has this main characteristics:

N	E	<k>	δ
23865	574744	48.17	0.002018

There are over 23 thousand users in the network and over 574 thousand edge. The average degree for a user is that of 48 edges, meaning that the expected number of relationships for a user is 48. The density of the network however is very small, which suggests that it is far from a fully connected network. Nevertheless, the are still working with a somewhat big network.

The degree distribution shown in figure 5 tells us that most of the nodes have a very low frequency, close to zero, which suggests that the mean computed is biased towards the nodes with low degree.

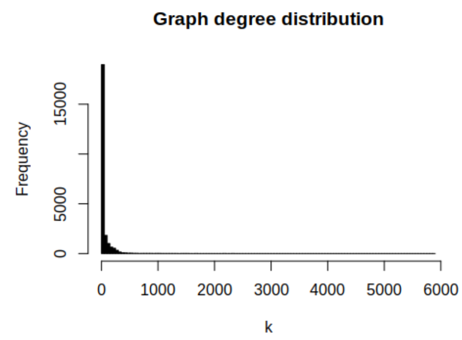

Figure 5: ”User-User graph degree distribution”

4 Simulation Approaches

For each model we are going to follow some general steps:

1) Create the user-user network G_{global}: the nodes being the users and the edge representing their relationship. If there exists an edge between two users then they are connected bidirectionally.
2) Let F be the set of n fake news, such that $F = \{f_0, f_1, .., f_n\}$.
3) For each $f \in F$, create a subgraph g of G_{global} such that $g \subseteq G_{global}$. This graph g will contain all the nodes that according to our dataset have shared the news plus the nodes 1 hop away from them that have not.
4) Find the β for each subgraph g.
5) Analyze the distribution of the β and try to extrapolate a β_f for a new network.

For each model, we have set a maximum time step $t_{max} = 48$, where each time step equal an hour having passed. This means that we are working assuming (1) the spread will die out at the 48th hour in the model and (2) that it has died out too at t_{max} in real life. Although we know this is not accurate we have settled on this assumption as explained in the Assumptions 3.2 section.

To we check whether the β_f for each model holds and compare the different approaches we are going to try a range of different β and for each one run 10 executions and get the mean number of

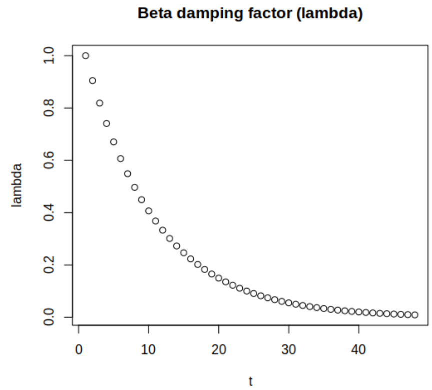

Figure 6: "λ damping factor for β in Baseline model"

nodes infected. Then, we are going to select the *beta* that has produced a model that after 48 time steps the number of nodes infected equals to those found in the news-user file. Once selected, we are going to test this β with the same 10 different news in the three models and we will see which one better fits them.

All models are an SI variation and are a graph $G = (V, E)$ in which each node i is in one of two states: infected or susceptible. A user is considered infected if and only if it has shared a fake news article.

4.1 Baseline Simulation

4.1.1 The Approach

The baseline will serve the purpose of being a model to which we can compare the others. This baseline works around two main characteristics: (1) we pick 3 random nodes on each fake new subgraph g. This means we can pick any node with the same probability regardless of their degree, centrality or any other metric. And (2), β decreases following an exponential function.

The worst case we are assuming here is that at the first time step, the fake news is very infectious and people is eager to read and share it, but as time passes this interest on the fake news quickly dies because the news gets old and people switch their interest onto other new fake news that may have appeared during this time.

4.1.2 The Model

There are two parameters: β, and t.

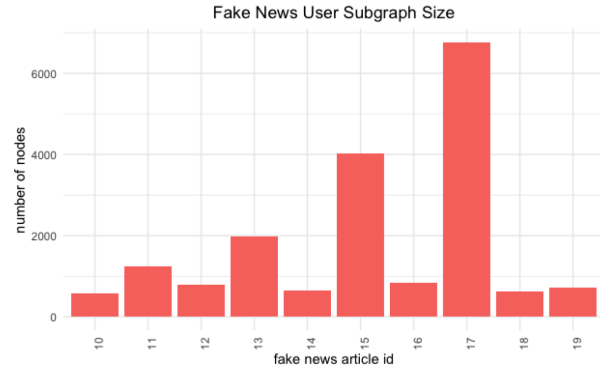

Figure 7

- β = [0,1]: Contributes to the infection probability.
- t: Number of time steps per simulation. Set to 48 by default.

As said in this model approach, β decreases following an exponential function defined as

$$f(t) = 1 - (0.1 * t)^e$$

which can be seen on figure 6. This damping factor will be referred to as λ and for each time step, the β_t on time step t will be defined as

$$\beta_t = \lambda * \beta = f(t) * \beta$$

4.1.3 The Results

Having β follow a decreasing exponential distribution may ensure that not everybody gets infected after 48 time steps if the starting β is small enough.

We have used $F_{train} = \{f_{10}, f_{11}, .., f_{19}\}$, so $|F_{train}| = 10$ as our training set to obtain the global β that better fits the models. The sizes (i.e. number of nodes) of the resulting sub-graphs for these fake news can be seen in figure 7.

Instead of using the average of the β obtained, we are weighting them according to how close the fitted values are to the real ones.

Let us define the ratio as follows:

$$r = \frac{min(\hat{y}, y)}{max(\hat{y}, y)}$$

And the weight related to a β such that:

$$w_\beta = \frac{r_\beta}{\sum_{i=0}^{10} r_i}$$

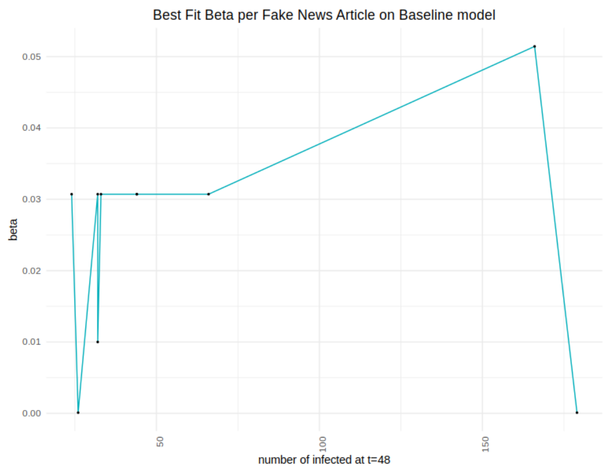

Figure 8: Each dot represents a separate fake news article.

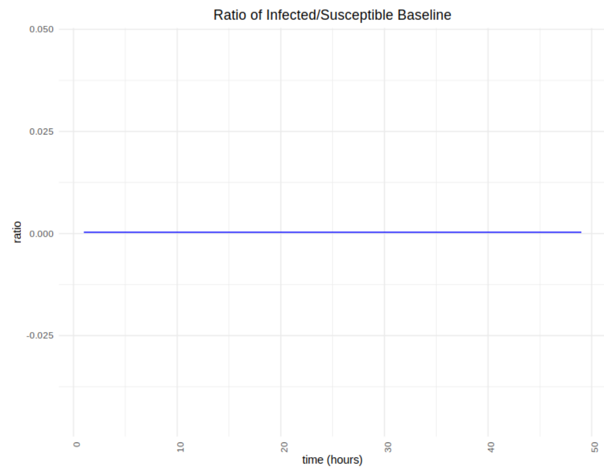

Figure 9: Baseline ratio is constant for different news

$$\sum_{i=0} r_i w_\beta = 1$$

Then, let us define the weighted βs formula β_w as:

$$\beta_w = \sum_{\beta i=0} w_{\beta i} * \beta \qquad (\beta_w)$$

In figure 8 we can see how for different sizes of the total infected nodes at $t = 48$, the beta changes. However, we cannot detect any correlation on the β fitted and the size of nodes infected after 48 timesteps.

So we get the average using the defined formula β_w of all these β and use this β_{avg} as the global β. Our result has been $\beta_{avg} = 0.03113859$.

However we can see on figure 9 how the ratio Infected/Susceptible is constant with different news. This is due to only the first 3 nodes getting infected and then, since β is so low, nobody else gets infected, so the ratio is constant as time progresses.

There are some cases where randomness plays its hand and the infection rate takes off and seems to adapt better. For example, for a fixed beta $\beta = 0.12$, which is the β found to best fit the fake news with id 4, we can see in figure 10 how it actually does not differ that much. However, with the average β found with all models, this evolution does not hold and the spread does not take off nor progress.

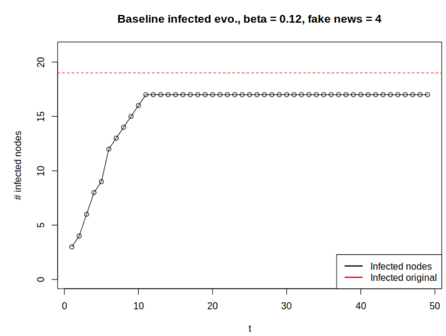

Figure 10: Evolution of fake news spread, in comparison to real final state

Finally, on figure 11 we can see how well our β fits different fake news ids and see whether they spread reaching a number of infected nodes similar to the original ones or not. As it is obvious, our model does not scale well and seems to either stale on the first nodes and not spread or spread very little, falling short in the predictions.

Because of this, we consider this baseline approach as a bad model but works as a starting point, nonetheless.

4.2 Dynamic β with peak and degree centrality.

4.2.1 The Approach

This model is based on the baseline but with some important changes: (1) we compute the degree centrality of each fake news subgraph g. (2) We have a variable p that selects how many nodes

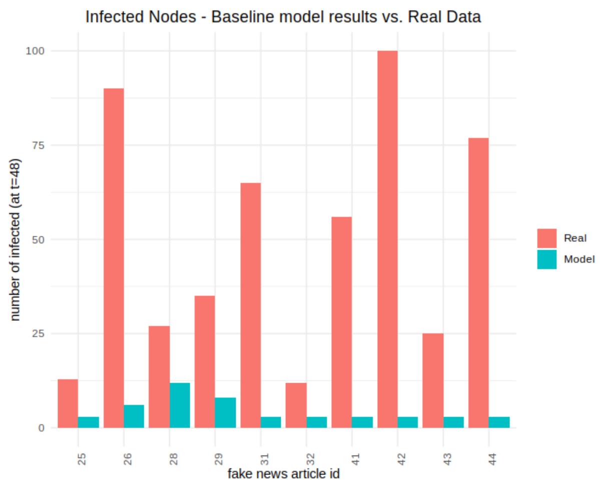

Figure 11: Baseline estimated values vs real ones

will be infected on $t = 1$. (3) We select the $\frac{p}{3}$ nodes with highest degree and infect them. The rest of $\frac{2p}{3}$ nodes are picked randomly.

The reasoning behind picking the top $\frac{p}{3}$ nodes that have greater degree centrality is because we can assume that those nodes that are well connected are so because they usually share news and content that may interest other people. Thus, they are also more prone to share willingly or not fake news than the average user.

And finally (4), β decreases following a Normal distribution, so again, a damping factor λ is used, such that $\lambda \sim N(\mu, \sigma)$, where $\mu = 3$ and $\sigma = 5$. This λ has been scaled as of to be in a range from 0 to 1. And so, we define $g(t) = \lambda * \beta$, which transforms how β progresses over time.

The evolution of λ can be seen on figure 12.

The reasoning behind this choice is that once some fake news are first posted on social media, their peak infectious state has not been reached yet, and it is not after some hours that it does so. Once the peak has been reached, it steadily decreases until dying out at time step 48.

4.2.2 The Model

Again, there are two parameters: β, and t, defined as before.

However, now β follows a Normal distribution which can be seen, again, on figure 12. This damping factor will be referred to as λ again and for each time step, the β_t on time step t will be defined as

$$\beta_t = \lambda * \beta = g(t) * \beta$$

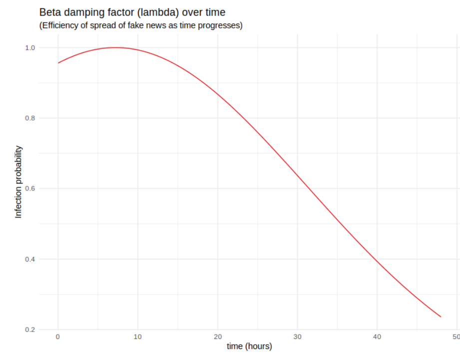

Figure 12: "λ damping factor for β in second model"

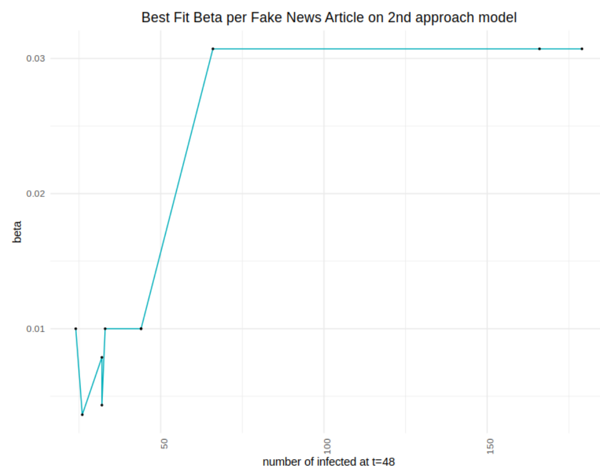

Figure 13: Each dot represents a separate fake news article.

The number of starting infected nodes, p, has been defaulted as 3 for all executions.

4.2.3 The Results

Having β follow the displaced Normal distribution seems to work better than the baseline by a far margin.

We have used again as with the Baseline $F_{test} = \{f_{10}, f_{11}, .., f_{19}\}$, so $|F_{test}| = 10$.

In figure 13 we can see how for different sizes of the total infected nodes at $t = 48$, the beta changes. We can notice that for our training fake news, once the number of users is greater than 70, the best fitting β does not change and adheres to a value a little over 0.03.

So as before we get the average of all these β using the weighted average defined at β_w and use this β_{avg} as the global β. Our result has been $\beta_{avg} = 0.01553287$.

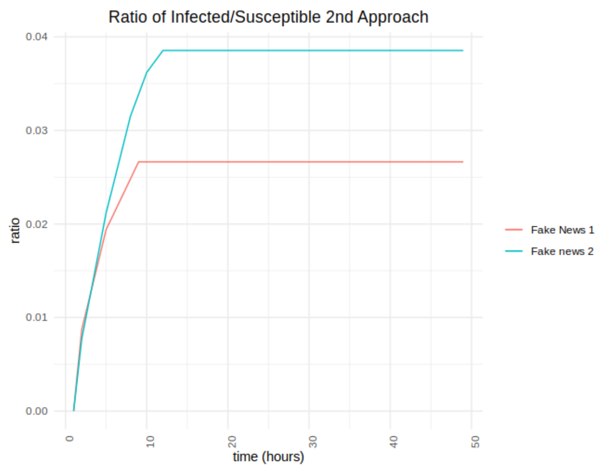

Figure 14: Baseline ratio is constant for different news

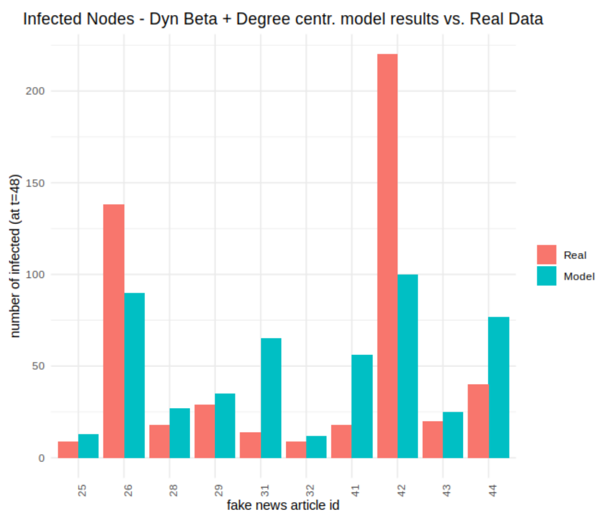

Figure 15: 2nd Approach estimated values vs real ones

However we can see on figure 14 how the ratio Infected/Susceptible is constant after the 15 iterations. This is due to the decay of β and the β_{avg} being so low.

In figure 15 we can see how well our β fits different fake news ids and see whether they spread reaching a number of infected nodes similar to the original ones or not. This model applies better to the fake news with a generalized β, although the results still fall short in most cases with the exception of the last fake news in the validation set, where it doubles the original number of nodes.

All in all, we consider this model as an improvement over the Baseline model, although we consider there is still margin for improvement.

4.3 Dynamic Beta with Multiple Parameters

4.3.1 The Approach

The simulation discussed in this section is a rough adaptation of the approach introduced in [Tambuscio et al., 2015]. There are three main ideas behind the approach: (1) Each fake news article has a different likelihood as being seen as truthful or fake by the general public, (2) The probability of a reader believing the fake news is heavily influenced by the number of their social connections who also believe the news and (3) Every person varies in the degree in which they fact check news articles. The idea behind (1) stems from the fact that fakes news articles vary wildly in their ability to convince readers they are true. For example, an article written by a well-spoken author that contains references to reputable news agencies (even if the references are false) will likely spread more quickly and aggressively than a poorly written news article by an author whose writing isn't elegant or contains spelling or grammatical errors. The idea behind (2) is that a reader is more likely to accept fake news as truthful if many of their social connections also believe the fake news, as that can be seen as corroborating evidence that the article is correct. The idea behind (3) follows the notion that some people are more likely to fact check a news article then others, thereby reducing their probability of believing any given article. This roughly translates to each user having a different beta in an SI model.

4.3.2 The Model

There are three parameters in this model: α, β, and t.

- α = [0,1]: The quality of a fake news article (i.e. how believable it is). Ranges from [0,1] with 0 indicating the article is obviously fake and 1 indicating the article is well written and is difficult to disprove.
- β = [0,1]: Contributes to the infection probability.
- t: Number of time steps per simulation.

The starting β is set at $t = 0$ but then becomes dynamic according to a damping factor that follows a normal distribution which can be see in figure 16. When a fake news article appears for the first time, very few people will have shared the article in the first few hours, then as time progresses more and more people will share the news and the rate of infection will climb. As more time passes the article will begin to be debunked and fewer people will share the news, decreasing the rate of infection. Finally, as more time passes the article will have run its duration and no more infections will occur. This process closely follows a normal distribution and can be easily adapted to differing time intervals. $\alpha = 0.5$ and $t = 48$ are used for all simulations. At each time step, the following equation, adapted from [Tambuscio et al., 2015], determines the probability of each susceptible node becoming infected.

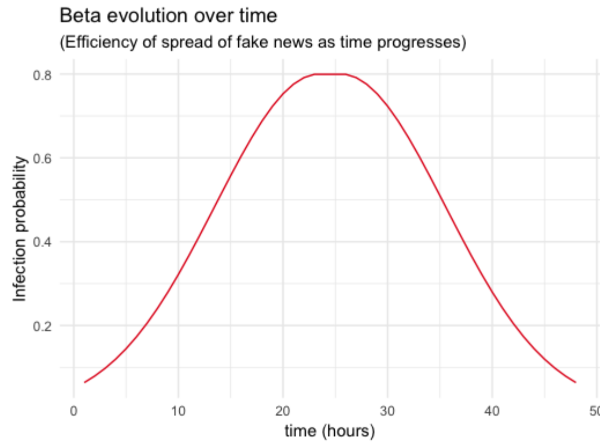

Figure 16: λ damping factor for β in third model

$$f_i(t) = \beta \frac{n_i^I(t) \cdot (1+\alpha)}{n_i^I(t) \cdot (1+\alpha) + n_i^F \cdot (1-\alpha)} \tag{1}$$

$f_i(t)$ is the infection probability of node i at time-step t. $n_i^I(t)$ is the number of neighbors of i (i.e. the number of social media connections) that are infected at time-step t. n_i^F is the number of neighbors that regularly fact check news articles. n_i^F is stochastic and is determined using a random normal distribution. Looking at equation 1 we can see that as α increases, the probability of infection increases as well. In contrast, as n_i^F increases the probability of infection decreases. This dynamic approach attempts to model the real-world scenario in which every consumer has a different probability of believing a fake news article depending on multiple closely related factors.

At the start of each simulation, the three nodes of G with the largest PageRank are infected. This method was chosen based on the idea that a social media user with many followers is likely to share news at a high rate to please their followers and is therefore more likely to share fake news.

4.3.3 The Results

In order to determine if this approach will provide a universal model that accurately simulates the spread of any fake news article over a network, we decided to find the best fit starting beta for 10 different fake news articles. We then plotted the beta against the total number of infected nodes at time-step $t = 48$, the final time-step in the simulation. The thought is that if we can find a single beta that works decently well for each of the 10 articles, or if the beta follows some sort of distinguishable distribution or pattern, then perhaps we can use the results to apply the model to future fake news.

To find the best fit beta for a fake news article, we ran the simulation multiple times starting with a very small beta and increasing the beta by a small amount each time. At the end of each simulation, we compared the total number of infected nodes in the network, to the actual number of infected nodes provided in the dataset. After all simulations were run, the beta with the closest number of infected nodes was chosen as best fit.

The results of this can be seen in figure 17 below. Each dot in the figure represents a distinct fake news article. The articles are ordered on the x-axis by number of infected nodes at the end of the simulation (t=48). We can see that, unfortunately, there is no noticeable correlation or pattern present. This leads us to believe that we likely can't find a starting beta that can be used universally for all fake news. However, there are still many aspects of the model that can be changed or tweaked, the details of which will be discussed briefly in the future work section.

Additionally, in order to see the rate of infection over time, the ratio of infected/susceptible was plotted for two fake news articles. This can be seen in figure 18. Although it is a bit unclear with such as small number of time-steps, we can see that the ratio does seem to follow the damping factor of the beta following a normal distribution. The infection rate is a bit slower at the beginning

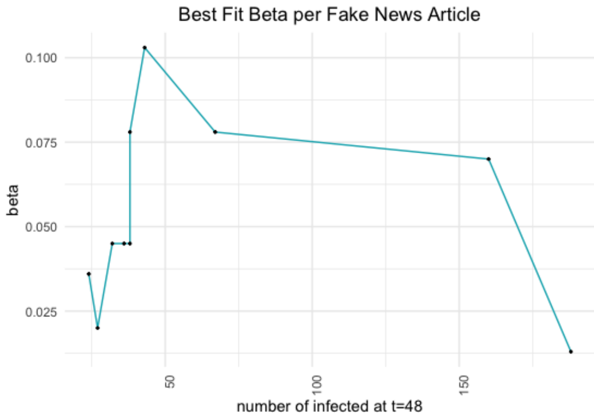

Figure 17: Each dot represents a separate fake news article.

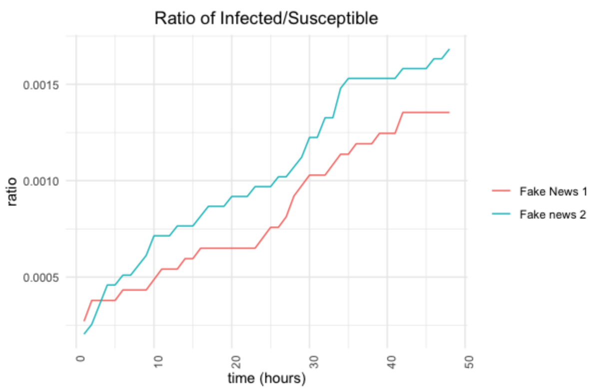

Figure 18

and end and a bit faster in the middle (t = 25-35).

Although there doesn't seem to be a universal beta that can be used to accurately simulate the spread of future fake news, it is worth testing with the results we have. For this, we have taken the weighted average beta from the 10 simulations and used this beta as the input parameter to the model. We then ran the model for 10 different fake news articles and compared the total number of infected at t=48 with the actual number of infected provided by the dataset. The average beta is $\beta_{average} = 0.0533$. The result of this test can be seen in figure 19.

5 Model Comparison

Now that we have seen how the models have performed individually, let's compare them.

To do so, we are using a custom metric that checks the quality of a model. Let y be the real number of infected nodes for a given fake news id,

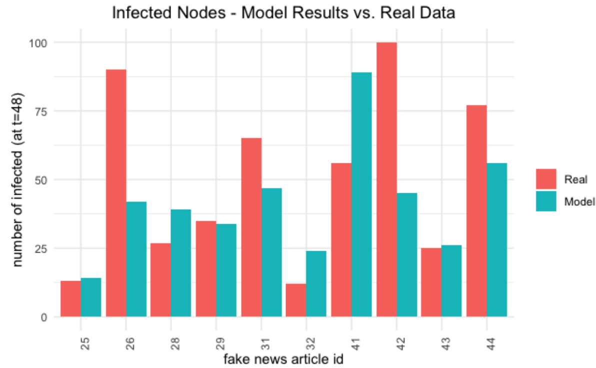

Figure 19: "Real" is the Actual number of infected nodes. "Model" is the number infected nodes resulting from the simulation.

$\hat{y}$ the infected nodes with the best beta found for the model and n the number of fake news we are validating with.

$$Q_{model} = \frac{\frac{1}{n}\sum_{i=1}^{n}(\hat{y} - y)^2}{var(y)}$$

The lower the value of Q, the better.

To validate the fitting of the beta we are doing so with fake news that we have not used before. Thus, we are treating these fake news as the validation set.

The validation set has been the same in all models.

The fake news are $F_{val} = \{f_{25}, f_{26}, f_{28}, f_{29}, f_{31}, f_{32}, f_{41}, f_{42}, f_{43}, f_{44}\}$, so $|F_{val}| = 10$.

These results can be seen on table 1, and we can see that the worst model is the baseline one. The dynamic β with degree centrality, the second model, way better than the baseline, so it seems that using a dynamic β following a displaced Normal distribution plus using degree centrality seems to do the trick.

And finally, the 3rd model, which is the one with a dynamic β with multiple parameters, is the one with the best performance. Looking back at figure 19 we can see that in many cases, the model infects roughly the same number of nodes as the real data. However, it doesn't perform so well with more infectious fake news.

We also tested the computation time (in seconds) for each of the models. Using the β_{avg} of each model, 10 simulations were run for each of the fake news in the validation set F_{val} and the

Model	Q
Baseline	2.9295
Dyn. β + degree centr.	1.26
Dyn. β + multiple params.	0.79

Table 1: Q metric for each model

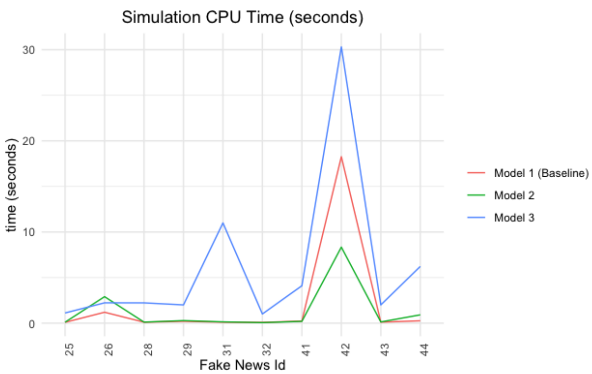

Figure 20: Simulation execution time for each model.

total CPU time was recorded and then averaged. The results of this can be seen in figure 20, and the specific average execution time is found in table 2.

Models' execution time (seconds)

News Id	Baseline	Dynamic β & <k>centr.	Dynamic β & mult. params
25	0.088	0.104	1.119
26	1.200	2.903	2.228
28	0.104	0.112	2.222
29	0.203	0.281	2.000
31	0.106	0.141	10.992
32	0.082	0.056	1.009
41	0.232	0.197	4.102
42	18.257	8.337	30.299
43	0.111	0.133	2.001
44	0.258	0.914	6.220

Table 2: Comparison of execution times for each model, same fake news

Additionally, the sub-graph size (i.e. number of nodes) for each fake news is shown in figure 21.

The first thing to notice is that model 3 (Dynamic β with multiple parameters) takes quite a bit longer than the other two models. For the most part, we can also see that larger sub-graphs take longer to simulate.

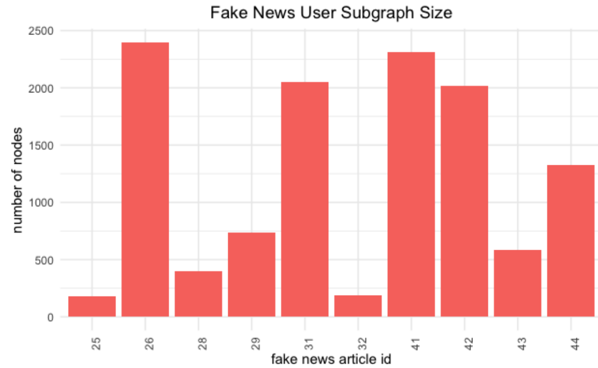

Figure 21

6 Conclusions and Future Work

The conclusions done from this project are many.

1) It is very difficult to find a Simulation system SI with a correct β that closely models how fake news spread over time. Because of this, even though the results have not been as satisfactory as we would have wanted them to, we believe it is a topic with a lot of potential for research.
2) In the same line of thought, we also believe that were we to have a dataset that models step by step how the news spread for different cases, we could have come up with a more sophisticated method that could model it.
3) Nevertheless, we are well aware that fake news play with people's emotions making the titles interesting, deceiving and thus *clickbaity*. By doing so, many fake news may spread with different infection rates, a result from taking into account when they may have been uploaded, how the title is worded or how many bots and which users, influential or not, have played their role on spreading them.
4) Even if a SI model that perfectly simulates the spread of fake news is found, it will be difficult to find a single set of parameters that can be applied in all cases. More likely, a good model will need to be continuously tweaked according to the fake news and associated network.

As for future work to do on this topic, we also have many ideas that we could not develop fully nor try because of the lack of time:

- Taking the worst case scenario with a very infective β, we have tried our luck trying to implement a mitigation system to try to stop the fake news spread. We have had no real luck on implementing this, for maybe the approach has not been the correct one. One may find an attempt of such implementation of this mitigation system on the function `src/l_approach /createFakeNewsSubgraphFromId`.
- Thus, try different approaches on how to mitigate the spread of fake news. From ranging to target the nodes with highest degree that have not yet been infected, to those that have a highest page rank or a higher betweenness centrality.
- As for modeling fake news, we could also try to take into account how many times a fake news has been shared to pinpoint those users as the possible source of the fake news spread. This would be applied instead of using degree centrality in the second model.
- A harder approach could also be taken, which would be to consider the graph as directed instead of undirected. Because of this, algorithms such as Page Rank may have a bigger impact on the results of selecting the source of the fake news or, in contraposition, selecting which nodes to immunize with *true news* so as to stop the spreading of the fake news.
- In the case of model 3, $\alpha = 0.5$ was used for all training and test simulations. It would be interesting to see how the model and β changes as α changes. It would also be interesting to use some sort of natural language process or pattern matching to classify fake news into various quality categories and then find a suitable α for each category.
- In the literature, there are already a large number of models proposed, it would be interesting to perform a more in-depth comparison of the models and their overall quality. This would then provide a reference for which types of models and approaches work best for fake news.

References

[Abbasi and Liu, 2013] Abbasi, M. A. and Liu, H. (2013). Measuring user credibility in social media. In *Social Computing, Behavioral-Cultural Modeling and Prediction - 6th International Conference, SBP 2013, Washington, DC, USA, April 2-5, 2013. Proceedings*, pages 441–448.

[Bessi et al., 2015] Bessi, A., Coletto, M., Davidescu, G. A., Scala, A., Caldarelli, G., and Quattrociocchi, W. (2015). Science vs conspiracy: Collective narratives in the age of misinformation. *PloS one*, 10(2):e0118093.

[Brede, 2012] Brede, M. (2012). *Networks - An Introduction.* mark e. j. newman. (2010, oxford university press.) \$65.38, £35.96 (hardcover), 772 pages. ISBN-978-0-19-920665-0. *Artificial Life*, 18(2):241–242.

[Castillo et al., 2014] Castillo, C., El-Haddad, M., Pfeffer, J., and Stempeck, M. (2014). Characterizing the life cycle of online news stories using social media reactions. In *Computer Supported Cooperative Work, CSCW '14, Baltimore, MD, USA, February 15-19, 2014*, pages 211–223.

[Del Vicario et al., 2016] Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., and Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559.

[Ecker et al., 2010] Ecker, U. K. H., Lewandowsky, S., and Tang, D. T. W. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, 38(8):1087–1100.

[Farajtabar et al., 2017] Farajtabar, M., Yang, J., Ye, X., Xu, H., Trivedi, R., Khalil, E. B., Li, S., Song, L., and Zha, H. (2017). Fake news mitigation via point process based intervention. *CoRR*, abs/1703.07823.

[Gupta et al., 2012] Gupta, M., Zhao, P., and Han, J. (2012). Evaluating event credibility on twitter. In *Proceedings of the Twelfth SIAM International Conference on Data Mining, Anaheim, California, USA, April 26-28, 2012.*, pages 153–164.

[Jin et al., 2014] Jin, Z., Cao, J., Jiang, Y., and Zhang, Y. (2014). News credibility evaluation on microblog with a hierarchical propagation model. In *2014 IEEE International Conference on Data Mining, ICDM 2014, Shenzhen, China, December 14-17, 2014*, pages 230–239.

[Pastor-Satorras and Vespignani, 2001] Pastor-Satorras, R. and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86:3200–3203.

[Saito et al., 2008] Saito, K., Nakano, R., and Kimura, M. (2008). Prediction of information diffusion probabilities for independent cascade model. In *Knowledge-Based Intelligent Information and Engineering Systems, 12th International Conference, KES 2008, Zagreb, Croatia, September 3-5, 2008, Proceedings, Part III*, pages 67–75.

[Shao et al., 2017] Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., and Menczer, F. (2017). The spread of fake news by social bots. *CoRR*, abs/1707.07592.

[Shu et al., 2018a] Shu, K., Bernard, H. R., and Liu, H. (2018a). Studying fake news via network analysis: Detection and mitigation. *CoRR*, abs/1804.10233.

[Shu et al., 2018b] Shu, K., Mahudeswaran, D., Wang, S., Lee, D., and Liu, H. (2018b). Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media. *CoRR*, abs/1809.01286.

[Shu et al., 2018c] Shu, K., Mahudeswaran, D., Wang, S., Lee, D., and Liu, H. (2018c). Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media. *arXiv preprint arXiv:1809.01286*.

[Shu et al., 2017a] Shu, K., Sliva, A., Wang, S., Tang, J., and Liu, H. (2017a). Fake news detection on social media: A data mining perspective. *CoRR*, abs/1708.01967.

[Shu et al., 2017b] Shu, K., Wang, S., and Liu, H. (2017b). Exploiting tri-relationship for fake news detection. *CoRR*, abs/1712.07709.

[Tambuscio et al., 2015] Tambuscio, M., Ruffo, G., Flammini, A., and Menczer, F. (2015). Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. pages 977–982.

[Vo and Lee, 2018] Vo, N. and Lee, K. (2018). The rise of guardians: Fact-checking URL recommendation to combat fake news. *CoRR*, abs/1806.07516.

[Vosoughi et al., 2018] Vosoughi, S., Roy, D., and Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380):1146–1151.

[Wang, 2017] Wang, W. Y. (2017). "liar, liar pants on fire": A new benchmark dataset for fake news detection. *CoRR*, abs/1705.00648.