CSE 6730 Modeling and Simulation

EPIDEMIOLOGICAL MODELING OF MISINFORMATION SPREAD DURING THE RUSSIA-UKRAINE WAR

Group 45:

GUNJAN GUPTA

OMAR JIMÉNEZ

VASISTHA SINGHAL

ABHIJEET TOMAR

https://github.gatech.edu/atomar39/cse6730-project-group45

EPIDEMIOLOGICAL MODELING OF MISINFORMATION SPREAD DURING THE RUSSIA-UKRAINE WAR [*]

Gunjan Gupta, Omar Jiménez, Vasistha Singhal, Abhijeet Tomar
College of Computing
Georgia Institute of Technology
Atlanta, Georgia
{ggupta68,ojimnez3,vvinod9,atomar39}@gatech.edu

ABSTRACT

Misinformation spread on social networks has emerged as an area of interest due to the increasing amount of information consumption from online sources. In this work we leverage the SEIZ model to describe misinformation spread on Twitter during the Russia-Ukraine War. We collected a dataset of 44927 tweets from March-April 2022 which mentioned the #IStandWithPutin hashtag – an element associated with the spread of misinformation related to this topic. Our results suggest that the adopted framework provides a reasonably accurate description of the system studied. In addition, we conduct simulations to gain insights on the potential dynamics of similar case studies related to misinformation spread on social networks. The use of mathematical modeling and simulations can potentially aid in characterizing the dynamics of misinformation spread on social networks and developing tools to target it and control its spread.

***K*eywords** Russia · Ukraine · Misinformation · Epidemiological Model · SEIZ Model · Twitter

1 Introduction

In recent years, social media has become one of the primary sources for diffusion of information. The highly interconnected nature of new social technologies provides users and entities with the ability to reach millions of people at little to no cost, which in turn can enable spread of misleading or inaccurate information. For example, in 2013, a false tweet claiming Barack Obama was injured in an explosion resulted in wiping out over $130 billion in stock value [1]. Similar events have also resulted in slowing relief efforts for hurricane disasters [2] and terrorist attacks [3]. Misinformation spread has also played an important role in other important events such as the 2016 presidential election [4]. To this end, the research community has been actively employing different strategies to better understand, target and simulate the spread of misinformation online. Among these include the development of crowdsourcing approaches [5], data-driven tools [6, 7], and simulations [8, 9, 10, 11, 12]. Community-based approaches like Birdwatch [5] allow users to identify content on Twitter they believe is misleading and write notes that provide informative context. On the other hand, data-driven tools typically leverage collected features (e.g. total number of followers/following, account creation date, replies, retweets) to build machine-learned models to predict whether a given piece of information is true or false. In this work we focus on how leveraging simulations helps us better understand misinformation spread online.

A standard approach that has been proven to be useful for simulating this problem is to draw analogies from mathematical epidemiological models and compare the dynamics of online information spread to those governing the spread of an infectious disease. Epidemiological models typically describe systems by designating their elements into corresponding *states*. For example, an element (e.g, a user on a social media platform) may become infected if it interacts with another infected user. These states may also evolve over time (e.g., a user may recover from a disease after a certain time), and these transitions are usually described by certain probabilities. Such models can thus help to predict the future course of an outbreak and to evaluate potential strategies to control an epidemic. In this work, we adopt an epidemiological approach to model misinformation spread on Twitter during the Russia-Ukraine War. Specifically, we propose to use

[*] *Citation*: Gupta, G. et al. "Epidemiological Modeling of Misinformation Spread During the Russia-Ukraine War"

the SEIZ model, which comprises states for Exposed (E) – users which require some time before becoming Infected, and Skeptics (Z) – users which have seen the misinformation item but decided not to engage in any reaction to it. This model has been previously shown to accurately capture the dynamics of information spread online [12, 13, 14].

The remainder of this paper is organized as follows. Section 2 discusses previous work that has been done related to the use of epidemiological modeling to study information spread within the social networks domain. In Section 3, we describe the data collection process and methods used to model misinformation spread. Section 4 discusses the main results and a detailed analysis of these observations. Finally, Section 5 concludes the paper with ideas for future work and a brief discussion of the limitations of our current approach.

2 Related Work

Epidemiological models have been previously employed to model the spread of information. In fact, research establishing comparisons between epidemics and rumours dates back to the last century [15]. Early works in the social media era used more primitive models like SIR (Susceptible, Infectious, Recovered) to study news spreading on Twitter [8], [9]. Xiong et al. [10] proposed a diffusion model containing four states: Susceptible, Contacted, Infected, and Refractory (SCIR) to characterize information propagation on online microblogs. Other approaches rely instead on the use of stochastic models to study information [11] and hoax [16] spread on social media. Tambuscio et al. also derived mean field approximations to estimate probability thresholds for model parameters.

More recently, the SEIZ (Susceptible, Exposed, Infected, Skeptical) model was proposed to model news and rumours on Twitter [12]. The key modification to this model is the incorporation of E and Z states, which correspond to users who knew about the item but decided not to engage in any interactions and those who heard about the news but needed some time before deciding how to respond, respectively. The SEIZ model has since been used to study misinformation spread during recent events such as the Black Lives Matter Movement [14] and COVID-19 [13]. Our work is motivated by the aforementioned studies and proposes to use the SEIZ model to study misinformation spread during the Russia-Ukraine War. To the best of our knowledge, this work is the first to explore this problem.

3 Methodology

3.1 Data Collection

We use *snscrape* to collect a dataset of 44927 tweets from March-April 2022 containing the #IStandWithPutin hashtag [17]. We note that similar datasets are usually collected within a smaller timeframe. However, our analysis required a longer time window to enable collecting sufficient data since it is known Twitter has been proactively banning accounts and removing tweets containing the hashtag of interest.

3.2 SEIZ Model

People can form differing, complicated beliefs when they are *Exposed* to items of misinformation on social media. Some people may have different viewpoints about the misinformation item; some others may need some time to come to believe it; or some others can be *Skeptical* to the accuracy of what they saw.

Based on the provided reasons, we decided to use the SEIZ model, which is a powerful model and is applicable to the spread of misinformation on Twitter because, as we mentioned before, it includes the *Skeptics* component and *Exposed* component which is more suitable for the process of spreading misinformation.

We define our SEIZ model conceptually [Fig. 1] having the following states and their interpretation wrt our modeling of misinformation:

- First, Susceptible (S) users are those who have not interacted with the content.
- Exposed (E) represents those who have seen it but had a delay of time before posting about it themselves.
- Infected (I) relates to users who have tweeted about it.
- Finally, Skeptic (Z) users are those who have seen the content but decided not to tweet about it.

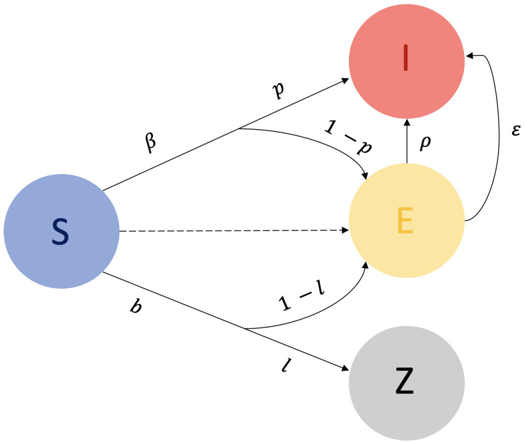


Figure 1: SEIZ Model showing different state and the transition probabilities

As part of using the SEIZ model, we try to encompass the possibility that it is possible for people to encounter a misinformation item on twitter and not execute any reaction. Furthermore, in the SEIZ model, individuals can be *Exposed* to a misinformation item but not tweet about it immediately.

We express our model mathematically using the following system of Ordinary Differential Equations (ODE):

$$\frac{d[S]}{dt} = -\beta S\frac{I}{N} - bS\frac{Z}{N} \tag{1}$$

$$\frac{d[E]}{dt} = (1-p)\beta S\frac{I}{N} + (1-l)\beta S\frac{Z}{N} - \rho E\frac{I}{N} - \epsilon E \tag{2}$$

$$\frac{d[I]}{dt} = p\beta S\frac{1}{N} + \rho E\frac{I}{N} + \epsilon E \tag{3}$$

$$\frac{d[Z]}{dt} = lbS\frac{Z}{N} \tag{4}$$

The various parameters used in the above set of ODEs are defined in the table below:

Parameter	Definition
β	Contact rate between S and I
b	Contact rate between S and Z
ρ	Contact rate between E and I
p	Probability of S to I given contact with I
1 - p	Probability of S to E given contact with I
ε	Transition rate of E to I
l	Probability of S to Z given contact with Z
1 - l	Probability of S to E given contact with Z

Table 1: Parameter Definitions for the SEIZ model

We start our simulation with a user that has not yet heard about the misinformation item and is in the state Susceptible (S). Upon coming in contact with an Infected (I) individual with a contact rate β, the user can theoretically instantly believe an item of misinformation with the probability of p, or that individual may have some doubts and need some time to analyze the misinformation item when they have time and move to the Exposed (E) state with a corresponding probability of (1-p).

We also have Skeptics (Z) users who have heard the misinformation but chosen to not tweet about it. These users transition from the Susceptible (S) state to Skeptics (Z) with rate b. These activities can cause two different possibilities. The first one is that it can cause turning the user into another Skeptic with the probability l. This means that the user decides not to tweet about the misinformation item maybe because they don't believe it or believe it but decide not to pass it on. The second possibility is that it can cause the inadvertent result of sending the user into the Exposed (E) compartment with the probability (1-l).

An Exposed (E) user can transition to Infected (I) state in two ways. The first one is that the people who are in the Exposed (E) compartment may come in contact with more Infected (I) individuals with a contact rate ρ and because of this contact they will become infected. Another possibility is that users in the Exposed (E) component can transfer to the Infected (I) component because of self-adoption with rate ε rather than because of having more contact with Infected (I) users.

3.3 Cellular Automaton Model

The SEIZ model discussed above assumes that each susceptible user is equally likely to come in contact with the infected or skeptical users. However, just like any social media network, users are more likely to see content from the users whom they are connected with in the network. In order to implement the transition of states depending on an individual users condition we implemented a cellular automaton model on a graph network.

3.3.1 Network

To run the cellular automaton simulations, we create two graphs - a Binomial or random graph and a Powerlaw Cluster graph. The random graph requires just one parameter to be set - the probability p for edge creation, whereas the Powerlaw Cluster graph requires 2 parameters - the number of random edges m to add to each new node in the graph and the probability p of adding a triangle after adding a random edge.

The number of nodes (N) in both the graphs is chosen as 2500 and they are randomly initialized to the S, I and Z states in the proportion $0.85 : 0.05 : 0.1$. Once the graph is set up and the nodes have been initialized, the SEIZ simulation is carried out for a fixed number of time steps (days in the plots). The simulation function takes several arguments and they are described below:

- t_1 - Threshold on the proportion of neighbors of a node that must be infected for it to be in contact with an infected node
- t_2 - Threshold on the proportion of neighbors of a node that must be skeptical for it to be in contact with a skeptical node
- p - The probability of transition from S to I given there is contact between S and I
- l - The probability of transition from S to Z given there is contact between S and Z
- k - The probability of transition from E to I given that at least one of the neighbors of an exposed node is infected.

3.3.2 Rules of CA model

Just like the ODE model, our CA model contains four states - Susceptible (S), Exposed (E), Infected (I) and Skeptics (Z). Initially the nodes can be in any of the 3 states - Susceptible (S), Infected (I), Skeptics (Z). The rules of state propagation are as follows -

1. If a node is in susceptible (S) state, and more than t_1 proportion of its neighbors are infected (I), then it can go to infected (I) state with probability p, or go to exposed (E) state with probability $1 - p$.
2. If a node is in the susceptible (S) state, and more than t_2 proportion of its neighbors are skeptics (Z), then it can go to skeptics (Z) state with probability l, or go to exposed (E) state with probability $1 - l$.
3. If a node is in exposed (E) state and one or more of its neighboring nodes is infected (I), then it can transition to infected (I) state with probability k.
4. If a node is in skeptics (Z) or infected (I) state it continues to remain in that state.

4 Experimental Results

4.1 SEIZ Model

4.1.1 Parameter Fitting

We fit our collected dataset of tweets containing the #IStandWithPutin hashtag with the Infected component of the SEIZ model. We use least squares method, which minimizes the L-2 norm.

$$e_{rel} = \frac{||I(t) - tweets(t)||_2}{||tweets(t)||_2} \tag{5}$$

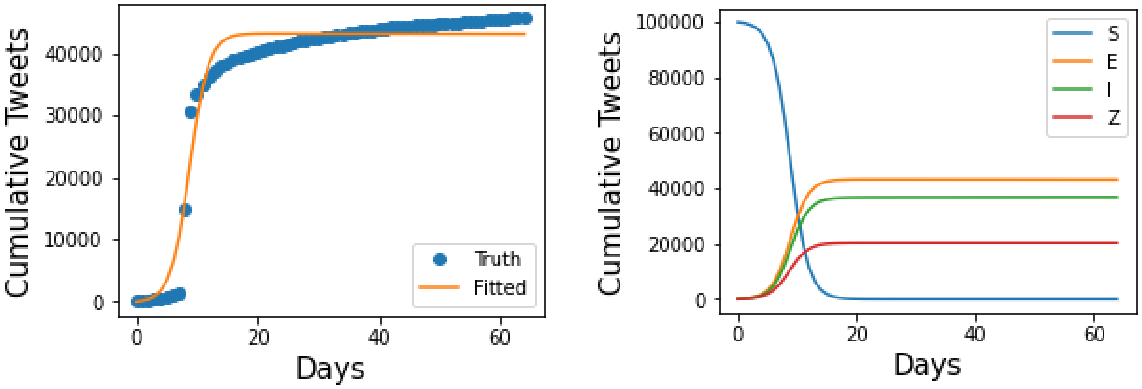
We use the coefficient of determination (R^2) as a measure to assess the adequacy of our model. R^2 is the proportion of variation in the data that is explained by the model. Mathematically, it is given by,

$$R^2 = \frac{SS_R}{SS_T} \tag{6}$$

The model parameters are sensitive to the initial conditions. We tried several combinations and the following initial conditions gave the best results. Assuming the total number of users to be $N = 100000$, we took $S_0 = 99771, I_0 = 29$ (from the data), $E_0 = 100, Z_0 = 100$. The set of optimal parameters that minimized this error are shown below in Table 2. For the model parameters obtained we see a high R^2 value of 96.9% suggesting the model explains the data reasonably well. The values of probability parameters p, l are $0.551, 0.602$ respectively, indicating that the users are slightly inclined towards the opinions of the users they come in contact with, whether it is an infected user or a skeptic. However, we see a reasonable chance of the user becoming exposed to the infectious tweet and then waiting for the information to be verified. In this case, we see a very low value of transition rate ϵ. We believe that this might be because of fact-checking from some users verifying that the information was not credible. Therefore the exposed users do not proactively participate in misinformation spread.

Table 2: Fitted parameter values

Parameter	Definition	Value
β	Contact rate between S and I	1.215
b	Contact rate between S and Z	1.001
ρ	Contact rate between E and I	1.000e-04
p	Probability of S to I given contact with I	0.551
ε	Transition rate of E to I	4.774e-52
l	Probability of S to Z given contact with Z	0.602



(a) Plot of infected users - Truth vs Fitted

(b) Simulation of user states using the fitted parameters

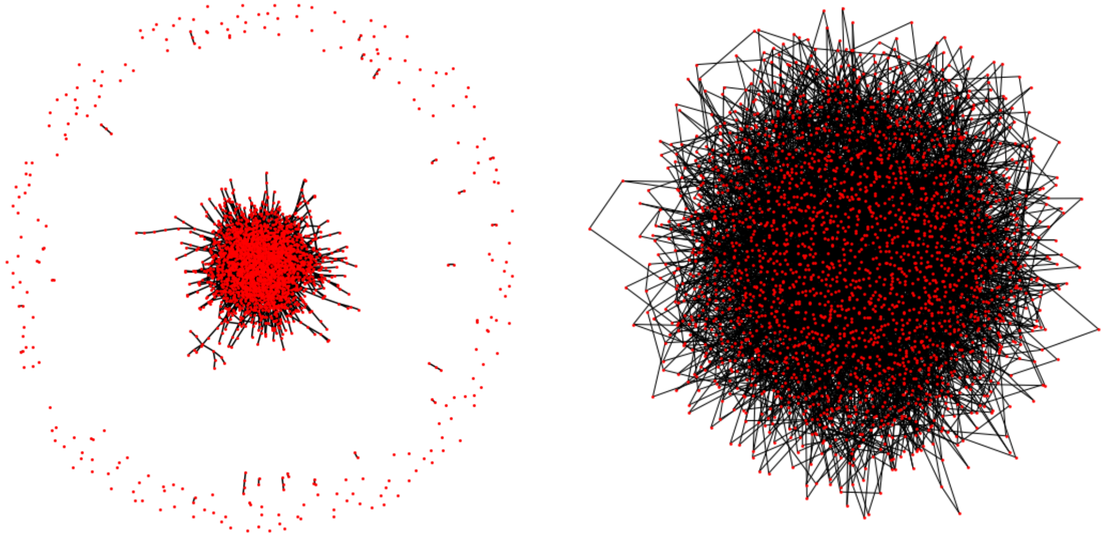
Figure 2: Different components of the SEIZ ODE model

4.2 Cellular Automaton Model

4.2.1 Simulations

Two sets of graphs are created. One set is very sparse, with $p = 0.001$ for the random graph and $m = 0.001N$ and $p = 0.01$ for the Powerlaw Cluster graph. The other set has $p = 0.005$ for the random graph and $m = 0.005N$ and $p = 0.01$ for the Powerlaw Cluster graph. We conducted simulations on these graphs with varying sparsity to see how the number of nodes in each state evolves. The arguments for the simulations were set as follows: $t_1 = 0.6, t_2 = 0.05, p = 0.75, l = 0.4, k = 0.05$. These values were chosen carefully to show the gradual evolution of all the states. To verify the sanity of our model, we also tested with extreme values of these parameters to see if the model would behave as expected in these obvious scenarios. For example, when t_1 and p were set to 1, all the susceptible nodes would transition to the infected state and the number of nodes in the skeptic and exposed states would remain the same throughout.

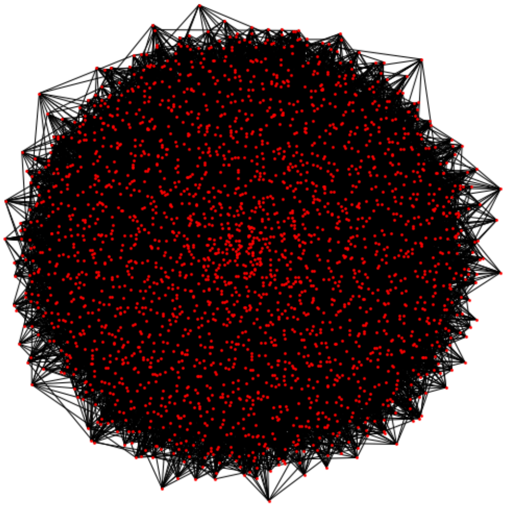
Figure 3: Visualizations of both the sparse graphs



(a) Visualization of the Random graph with 2500 nodes and p = 0.001

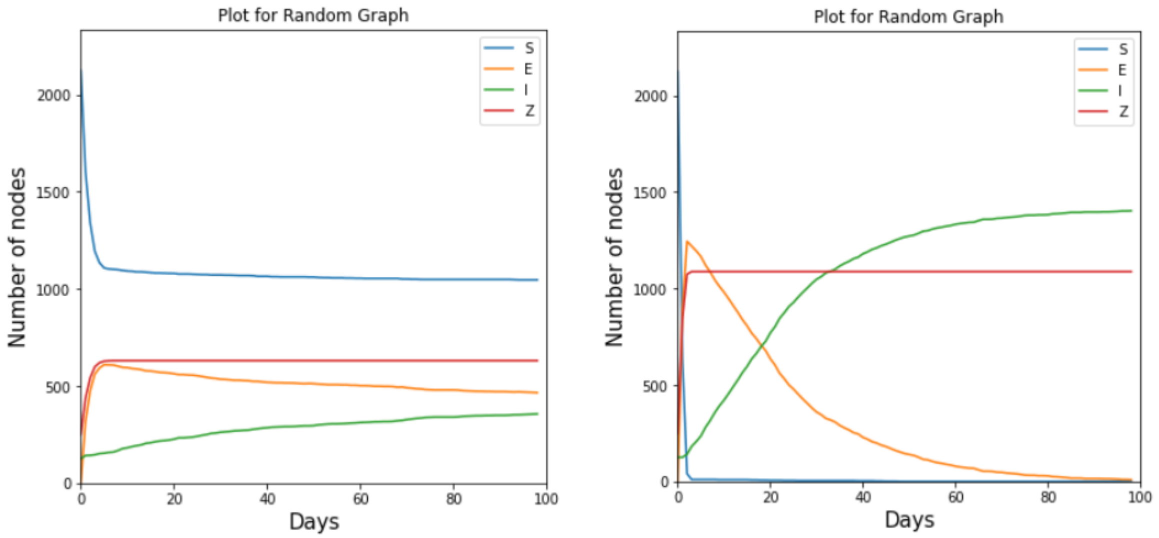
(b) Visualization of the Powerlaw Cluster graph with 2500 nodes and m = 0.001*N

Figure 4: Visualization of the dense graph (random and powerlaw cluster graphs look similar in the dense setting)



From Fig.3, we see the difference in the structures of the random and Powerlaw Cluster graphs. Despite having sparse connections (low m), the Powerlaw Cluster graphs seems to be well connected when compared to the random graph which has lot of isolated nodes and small clusters with few nodes. The consequence of this is clearly visible from the stark difference in the behavior of the system from the figures 5a and 6a. For the random graph, the majority of the nodes remain in the susceptible state because the contact with the nodes in the other states is minimal due to the disconnected nature of the network. Whereas, for the powerlaw cluster graph, the majority of nodes have become infected because the network is still well connected due to the presence of clusters.

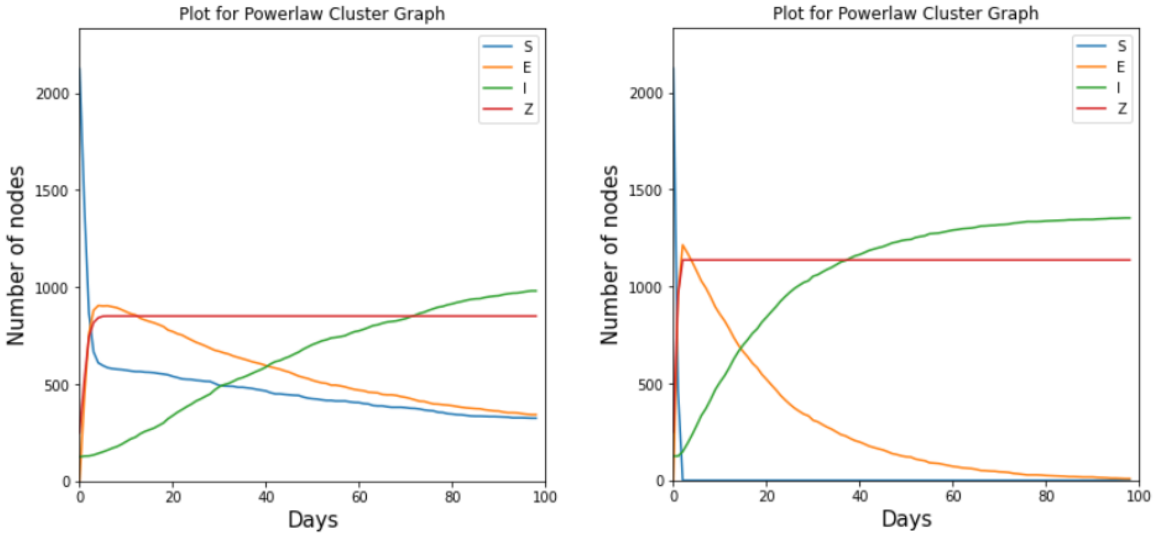
Figure 5: Simulation results for the random graph



(a) Simulation on a Random graph with 2500 nodes and p = 0.001

(b) Simulation on a Random graph with 2500 nodes and p = 0.005

Figure 6: Simulation results for the Powerlaw Cluster graph



(a) Simulation on a Powerlaw Cluster graph with 2500 nodes and m=0.001*N

(b) Simulation on a Powerlaw Cluster graph with 2500 nodes and m=0.005*N

From the figures 5 and 6 We see that the plots are drastically different when we vary the sparsity of the graph. For a very sparse graph (with p=0.001), the susceptible population remains the majority because its interactions with the infected population are very few. So, the infected population increases very slowly. The same trend is also observed for the skeptic population.

On the other hand, for the graphs with higher density, the susceptible population is rapidly converted to either the infected or the skeptic population. In the beginning, we see that the population in the exposed state rises but then the exposed population gets converted to the infected population after some lag. This is in accordance with the definition of the E state in the ODE model as well.

5 Discussion and Conclusions

Conclusions: In this project, we developed models to study how misinformation concerning the Ukraine-Russia war was propagated on Twitter. We assume the SEIZ epidemiology model for misinformation spread and use the tweet data acquired from the web to obtain optimal model parameters. We achieved an $R^2 = 0.969$ for the model indicating that the model explains the data adequately. We also study the spread of misinformation spread on dense and sparse graphs using a Cellular Automaton Model which was implemented using the *networkx* library in Python. The results from this study can provide strategies for slowing down the spread of misinformation. For instance, if there is a way to impose sparsity in the twitter network by removing tweets involving misinformation and/or banning users who share such tweets, then we can slow down the spread drastically, as seen in Fig. 5a. We see that even with a network that is not very dense ($p = 0.005$, Fig. 5b), the misinformation happens to spread very quickly and affect the entire susceptible population in a short period of time.

Limitations and Future Work: The models we used have certain limitations. In our SEIZ model, we are trying to capture time sensitive phenomena and for our experiments, we were limited to capturing tweets data on a per day basis. A finer granularity on the scale of hours or minutes and having a highly reliable dataset that contains the tweets and hashtags labeled as misinformation could possibly help to model the spread better. We also suggest finding data that has enough information to be able to infer the interconnectedness of the network from Twitter user data (user's followers, user's retweets, etc.) to run the cellular automaton simulations. Applying these models to datasets from other social media platforms such as YouTube, Facebook, etc. might reveal interesting insights about human behaviour on social media platforms. This knowledge can be crucial for developing strategies to mitigate the spread of misinformation on social media platforms.

References

[1] Kenneth Rapoza. Can 'fake news' impact the stock market?, Jun 2021.

[2] Aditi Gupta, Hemank Lamba, Ponnurangam Kumaraguru, and Anupam Joshi. Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In *Proceedings of the 22nd international conference on World Wide Web*, pages 729–736, 2013.

[3] John Woodrow Cox Marc Fisher and Peter Hermann. Pizzagate: From rumor, to hashtag, to gunfire, Dec 2016.

[4] Andrew Guess, Brendan Nyhan, and Jason Reifler. Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 us presidential campaign. *European Research Council*, 9(3):4, 2018.

[5] Twitter. Birdwatch: https://twitter.github.io/birdwatch/.

[6] Stefan Helmstetter and Heiko Paulheim. Weakly supervised learning for fake news detection on twitter. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 274–277. IEEE, 2018.

[7] Cody Buntain and Jennifer Golbeck. Automatically identifying fake news in popular twitter threads. In *2017 IEEE International Conference on Smart Cloud (SmartCloud)*, pages 208–215. IEEE, 2017.

[8] Saeed Abdullah and Xindong Wu. An epidemic model for news spreading on twitter. In *2011 IEEE 23rd international conference on tools with artificial intelligence*, pages 163–169. IEEE, 2011.

[9] Hui Wang, Lin Deng, Yi-Shuan Huang, and Shu Zhao. A variant epidemic propogation model suitable for rumor spreading in online social network. In *2012 International Conference on Machine Learning and Cybernetics*, volume 4, pages 1258–1262. IEEE, 2012.

[10] Fei Xiong, Yun Liu, Zhen-jiang Zhang, Jiang Zhu, and Ying Zhang. An information diffusion model based on retweeting mechanism for online social media. *Physics Letters A*, 376(30-31):2103–2108, 2012.

[11] Jun-Jun Cheng, Yun Liu, Bo Shen, and Wei-Guo Yuan. An epidemic model of rumor diffusion in online social networks. *The European Physical Journal B*, 86(1):1–7, 2013.

[12] Fang Jin, Edward Dougherty, Parang Saraf, Yang Cao, and Naren Ramakrishnan. Epidemiological modeling of news and rumors on twitter. In *Proceedings of the 7th workshop on social network mining and analysis*, pages 1–9, 2013.

[13] Maryam Maleki, Mohammad Arani, Erik Buchholz, Esther Mead, and Nitin Agarwal. Applying an epidemiological model to evaluate the propagation of misinformation and legitimate covid-19-related information on twitter. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 23–34. Springer, 2021.

[14] Maryam Maleki, Esther Mead, Mohammad Arani, and Nitin Agarwal. Using an epidemiological model to study the spread of misinformation during the black lives matter movement. *arXiv preprint arXiv:2103.12191*, 2021.

[15] Daryl J Daley and David G Kendall. Epidemics and rumours. *Nature*, 204(4963):1118–1118, 1964.

[16] Marcella Tambuscio, Giancarlo Ruffo, Alessandro Flammini, and Filippo Menczer. Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. In *Proceedings of the 24th international conference on World Wide Web*, pages 977–982, 2015.

[17] Ben Collins and Natasha Korecki. Twitter bans over 100 accounts that pushed istandwithputin, Mar 2022.

6 Appendix: Division of Labor

- Vasistha - Cellular automaton implementation, CA experiments and analysis for dense and sparse graph networks
- Gunjan - SEIZ model fitting and parameters estimation, Experiment analysis of SEIZ model and CA models
- Omar - Data scraping and cleaning, Contemporary literature and related works review, SEIZ modeling
- Abhijeet - SEIZ model setup using appropriate data structures for parameter estimation, Visualisations for the experiment results