

# Dynamic SEIZ in Online Social Networks: Epidemiological Modeling of Untrue Information

Akanksha Mathur<sup>1</sup>, Chandra Prakash Gupta<sup>2</sup>  
Computer Science and Engineering Department  
Rajasthan Technical University  
Kota

**Abstract**—The epidemic propagation of untrue information in online social networks leads to potential damage to society. This phenomenon has attracted attention to researchers on a faster spread of false information. Epidemic models such as SI, SIS, SIR, developed to study the infection spread on social media. This paper uses SEIZ, an enhanced epidemic model classifies the overall population in four classes (i.e. Susceptible, Exposed, Infected, Skeptic). It uses probabilities of transition from one state to another state to characterize misinformation from actual information. It suffers from two limitations i.e. the rate of change of population and state transition probabilities considered constant for the entire period of observation. In this paper, a dynamic SEIZ computes the rate of change of population at fixed intervals and the predictions based on the new rates periodically. Research findings on Twitter data have indicated that this model gives more accuracy by early indications of being untrue information.

**Keywords**—Information diffusion; epidemic models; SEIZ; rumor detection

## I. INTRODUCTION

The term “Social Networks” (SNs) was first coined Barnes in 1954[1]. It emerged in the form of e-mail and now used in many applications. In the last few years, SNs also referred to as social media have spread at a phenomenal pace covering people in every walk of human society. It has allowed people to share ideas, opinions, and seek information quickly and effectively. Emergence and boom of social networking websites like Facebook, Twitter and Reddit have proved useful in catastrophe control required in situations like flood, storm, and earthquake. These platforms have been equally effective in mitigating situations caused due to terrorist attacks, shootouts, and other similar situations. All these platforms are open platforms on which people can share whatever they wish with almost no censorship or filtering or verification of the contents by the media operators.

The enormous use of online social networks (OSNs) by millions of daily active users leads to radical changes in information sharing and communication among users as in [2]. Statistics show that over the years 2016-2018, about 20% of adults received news from U.S. social media websites [3]. According to study results [4], in 2017, the news on selected countries named Television as their source of news with 41 %, while 44% identified the internet.

(Including social media) as their source of news, as per in Fig. 1.

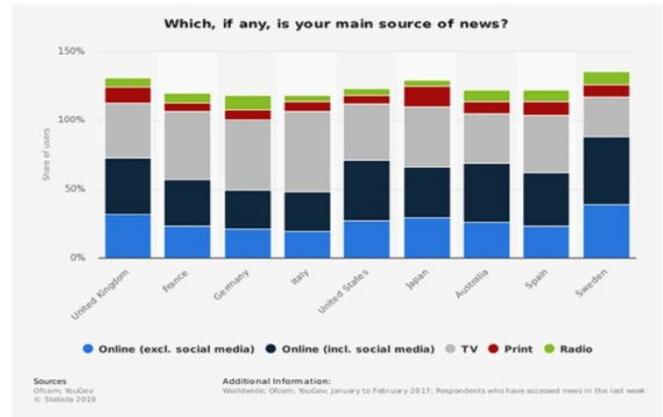


Fig. 1. Selected Country News Sources (2017) [5].

Websites intentionally publish hoaxes, rumors and misleading information popping up all over the web, and are shared on social media to extend their reach [5]. Terms such as "false news", "post-truth" and alternative facts" that is strongly associated with exposure of news in media. For example, stories such as Hillary Clinton's commercialism weapons to ISIS and Pope Francis support Trump for President liked and commented thousands of times on Facebook. Also, 14% of the people admitted that they have deliberately shared a fake political news story [6]. It is clear from this that the false news will continue to gain attention as long as users are still willing to share online. Even the current online checking system, such as Factcheck.org and Politifact.com, are standard detection approaches supported by professionals who specialize in the verification of the story item. Besides, every day a large amount of real-time information posted, answered, replied and shared via online social networks, which makes the detection of false news even harder.

In recent years, extensive research carried out on the establishment of an automatic and effective system or framework for the detection of online fake news and thus retrieving the useful and valuable information. Identifying legitimate information from a large amount of data available at different platforms is a challenging task because of the heterogeneous and dynamic nature of online social communications.

Misinformation can be defined as any malicious information that has spread deliberately or unintentionally (i.e.

without realizing it is untrue) [7]. Dissemination of false information has currently become identical with fake news. Moreover, the longer misleading information spreads within the network of a trusted source, the greater the impact. However, it is desirable to avoid unnecessary propagation of messages that causes an excessive burden on infrastructure that is used as long as-over emergencies [8].

The problem of false information or rumors in online social networks usually observed as misinformation or rumor blocking [9]. In addition, the lowest set of users can minimize the effect of spreading achieved by the originators of rumor. Other rumors distinguished by the subset of users act as the defender against the dissemination of misleading information sources [7]. The goal was to find the smallest subset of influential users and restrict the spread of misinformation to the specified ratio in a given set of the population. The proliferation of competitive campaigns evolving and aiming towards the selection of the appropriate subset of nodes such as the influence of misinformation campaigns also decreased as in [10].

For the effective and efficient solution to the problem, a series of research questions that need to address-

- How to effectively eradicate misinformation and understand the news propagation ways
- How contradictory information affects an individual's decision to accept and share the contents
- How can misinformation sources are early identified
- How to visualize and classify the misinformation of high dimensional real-time data

#### A. Research Challenges and Issues

The research challenges from the motivation of various papers on online fake news detection and some missing research directions are as follows:

- Determining an effective and reliable detection system for the analysis of rumor content in OSN's as early as possible.
- The studies suggest that online detection of fake content diversified in terms of methodologies, objectives, and research areas. Summarization of varied techniques and methods, with representative features, and evaluation of existing detection systems is required.
- The binary classification is not sufficient to classify the characteristics of deceptive information because of the complexity of the detection of untrue information. Data visualization is a powerful tool to illustrate the heterogeneity of social structures and the dissemination of social information online modes.
- In addition to the detection system, two aspects of promising research aimed at combating the false news. They are: (a) the early prediction of false news (i.e. it is very significant to identify trends or potential false news as early as possible.); and (b) False news

intervention (i.e. from the historical data, the initial predictions appear false news or rumors before the intervention). This work tries to assess whether the system is based on online communication information is news or rumor. However, it is extremely significant to identify trends or potential false news as early as possible.

#### B. Contributions

The contribution of the work summarized as follows:

- The SEIZ epidemic model used to simulate the Twitter dataset in real-time. The non-linear least square optimization used over tweet data to underlying network ODEs and demonstrates how false information modeled with the SEIZ.
- Comparison and data analysis between classical and dynamic SEIZ in the network provides an early indication and distinguishes the rate of change of population with time on Twitter data, respectively.
- The parameters-Exposed Ratio (RSI) and Reproductive Number (ROIZ) that show the capability of the SEIZ model to quantify information to be a rumor in compartment transition dynamics. In addition, these measures provide suggestive results for the rate-infected user affects the exposed compartment and susceptible compartments.

The paper structure is as follows; Section 2 presents a comprehensive description of existing work in epidemic models and rumor models. Next, Section 3 demonstrates a detailed analysis of SEIZ with experimental evaluation by fitting it with the dataset. Section 4 provides a comparative study with a classical and modified dynamic SEIZ model with results. It also determines early rumor detection ratios for identifying them in the network. Finally, Section 5 concludes the research study.

## II. RELATED WORK

A significant amount of work studied and modeled into research on information diffusion in social networks, which includes epidemic models and rumor modeling techniques in literature with their significance. Below is the summarization of these techniques from the literature.

#### A. Epidemic Models

Diffusion of information similar to the spread of the epidemic, but there are differences. Information diffusion is related to time, the strength of the relationship, information content, social factors, network structure, etc. Researchers have made ongoing improvements based on the classical model, developing new models such SI model (i.e. Susceptible, Infected) by [11], SIS model (i.e. Susceptible, Infected, Susceptible) by [12], SIR model (i.e. Susceptible, Infected, Recovered) by [13], SIRS model (i.e. Susceptible, Infected, recovered, Susceptible) by [14]. The SEIZ model (i.e. Susceptible, Exposed, Infected, and Skeptic) proposed by [15] takes a distinct approach by introducing a state Exposed(E). It is quite an effective mapping disease model with rumor models [16]. Several papers study the properties of

spreading information online social networks with emphasis on Twitter [17][18](Wang et al., 2013).

Information diffusion includes a set of characteristics summarized as follows: Diffusion minimization problem which identifies a small set of individuals with whom the dissemination of information diffusion is higher [19][14]. The role of users before the spread and classifying them consistent with their activity (i.e. re-tweeting, generating new tweets, commenting tweets, mentioning, replying, no. of followers so on) [13][15][17]. The re-tweet chain and the range of information diffusion [17], Hash-tags[20], temporal and topical patterns [21]. Modeling and studying the dynamics of information diffusion through the model of temporal dynamics of the epidemic and the diffusion of information on Twitter by [22].

The SEIR model was developed by adding another compartment Exposed (E) nodes supported the SIR model [23]. It is a dynamic information propagation method to analyze the impact of the frequency of the logged-in user and various friends. The results showed that the frequency of the logged-in user is directly proportional to the speed and range of transmission of information. S-SEIR model designed for single-layer social networks that support SEIR and its transmission of information depends on user activities [24]. SCIR model assumes that all followers assigned to the contact state 'C' when the user posts a message. It is a model for Micro-blogs by adding a contacted state (C) and provides an online topic spreading model [23][24]. In addition, the state of the followers may change according to a certain probability, the change in transmission or immune users after a particular duration. In [25], proposed model IR-SIR based on the SIR model simulating the adoption and abandonment of user views by adding a recovery process of the mechanics of the infection and verified the process on Facebook and Google. In [26], proposed the FSIR model that considers the effect of neighbors on an individual within the diffusion information. The threshold for spreading of information is zero and affects information diffusion. ESIS model based on SIS by [27], where the information transmitted between people has expressed a sort of emotional information.

Compartmental models are a mathematical approach applied to measure and predict the spread of various infectious diseases. The method of misinformation diffusion is usually a similar approach as a virus spreading process. In transmission epidemics, there is each user infected with viruses and can become susceptible to viruses. In [15], proposed the SEIZ model where the skeptics are the individuals becoming immune to infection. Although it is similar to the removed (R) individuals, skeptic transitions directly from the susceptible state and their interaction will still affect different compartments as well.

### B. Rumor Modeling

In 1964, analogy between the spreading of infectious disease transition and information propagation in the network was given [28]. In particular, it emphasizes on a mathematical model for the spreading of infection in several ways, depending on a mechanism to describe the growth and decay process. Later experimentation conducted that guaranteed

greedy approximate solution applied to Facebook [10]. According to the results, minimizing the spread of misinformation in social networks (SNs) is an NP-hard problem. Similarly, rumors detection problems in the micro-blog explore the effectiveness of into three categories (i.e. content-based, network-based, and micro-blog memes) for correctly identifying rumors [14]. [29]analyzed information credibility of news propagated through Twitter. The results are evidence that there are measurable differences in the way messages propagate, with a range of 70% to 80% of precision and accuracy rates. [30] Modeled rumor propagation in the network with the model Susceptible- Infected (SI) and constructed an estimator for the rumor source. They established a maximum likelihood (ML) estimator for a graphic class to find rumor sources in the network. The stochastic epidemic study on infection spread and rumors on networks, focusing on the SIR epidemic by ignoring density correlation between neighboring nodes investigated by [31]. This method computed mathematically and compared with the mean-field deterministic model over the set of equations.

### C. SEIZ Model

Compartmental models are a mathematical approach used to evaluate and predict the spread of various infectious diseases [35]. SEIZ model is a compartmental model that breaks the population into distinct compartments and establishing parameters for the rates at which the population transitions between compartments. These parameters obtained by looking at the relationships between each class of the population and making assumptions about the disease. With this, a set of differential equations generated to make predictions about the spread of the disease. A basic approach that is more practical than the SI model. Here the population divided into three compartments: Susceptible (S), who are the individuals at risk to the disease; Infected (I) which are those who have the disease and are capable of transmitting it. Removed I which are individuals who can no longer be infected or infect anyone else, so they either died from the disease or they have recovered and are now immune [32]. A great application of this model would be to a disease like measles, where once recovered, an individual has immunity for life. One of the assumptions made in this Susceptible-Infected-Removed, or SIR, model, is that the overall population,  $N$ , and is constant.

$$\frac{dN}{dt} = 0, \tag{1}$$

$$N = S + I + R. \tag{2}$$

Suppose that individuals only transition from  $S$  to  $I$  at a contact rate of  $r > 0$  and from  $I$  to  $R$  at a transition rate of  $a > 0$ , as shown in Fig. 2.

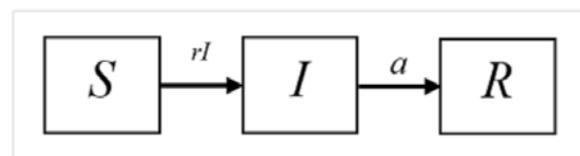


Fig. 2. Transition Rates of SIR Model.

These relationships then used to generate a system of ODEs:

$$\frac{dS}{dt} = -rS\left(\frac{I}{N}\right) \tag{3}$$

$$\frac{dI}{dt} = rS\left(\frac{I}{N}\right) - aI \tag{4}$$

$$\frac{dR}{dt} = aI \tag{5}$$

The rate at which the individual SI transition depends on  $r$  as well as the number of individuals infected. The rate of change depends on Infected (I) because if someone is in contact with an infected individual and becomes infected. The transition rate from I to R, however, is only dependent on ‘a’ as per the equations 3, 4 and 5. The transition for each individual out of the infected compartment is independent of everyone else for this model. Another simple model used is the Susceptible-Infected-Susceptible, or SIS, model as shown in Fig. 3.

In this case, it has assumed that instead of dying or becoming immune, the infected individual transition back into the susceptible class. Thus, this model used for diseases such as gonorrhea, because individual does not gain immunity after recovering from an infection. Though this model is not perfect, its simplicity can give us a good idea of what is going on. If we assume that individuals transition from S to I with a contact rate of  $r > 0$  and back from I to S with a transition rate of  $\lambda > 0$ , gives.

With these relationships, the following system of ODEs generated:

$$\frac{dS}{dt} = -rS\left(\frac{I}{N}\right) + \lambda I \tag{6}$$

$$\frac{dI}{dt} = rS\left(\frac{I}{N}\right) - \lambda I \tag{7}$$

According to the SIR model, the rate of transition from (I) is independent of anything except the parameter ‘ $\lambda$ ’. In addition, the rate of transition from S to I depends on the parameter ‘ $r$ ’, and the number of people in the infected compartment as per the equation 6 and 7.

One major drawback of SIS and SIR models is that a susceptible individual moves to the infected class immediately after exposure to an infected individual. This is usually not the case; however, as many pathogens take time to develop during a phase of incubation. For these instances, an Exposed Compartment (E) incorporated to indicate those who have met an infected individual, but it takes some time in the model to infect them themselves. From this extension, the SEIZ model explored one more compartment Skeptics (Z). Z recruit from the susceptible population with rate ‘ $b$ ’, but results either in

transforming the individual into another skeptic (Z) with probability (l), or can have unintended effect of sending that individual to the incubator class with probability (1-l).

In this model, the susceptibility immediately infected with probability ‘ $p$ ’, and ‘(1-p)’ is the possibility of an individual transiting to the incubator class instead, from which they adopted. ‘ $N(t)$ ’ denotes the total population where the network has a disease-free status with  $S^*=N$ ,  $E^*=I^*=Z^*=0$ . Fig. 4 illustrates the relationship between each compartment.

Table I and Table II provide description of each parameter of this model. This provides a more intuitive look into the model and relates the relationships above with the actual equations:

With the relationships between each compartment described by the parameters above, we have the following set of ODEs:

$$\frac{ds}{dt} = -\beta * S\left(\frac{I}{N}\right) - b * S\left(\frac{Z}{N}\right) \tag{8}$$

$$\frac{dE}{dt} = (1-p)\beta S\left(\frac{I}{N}\right) + (1-l)bS\left(\frac{Z}{N}\right) - \rho E\left(\frac{1}{N}\right) - \epsilon E \tag{9}$$

$$\frac{dI}{dt} = p\beta S\left(\frac{I}{N}\right) + \rho E\left(\frac{I}{N}\right) + \epsilon E \tag{10}$$

$$\frac{dZ}{dt} = lbS\left(\frac{Z}{N}\right) \tag{11}$$

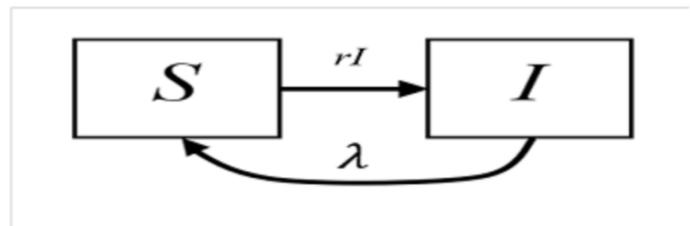


Fig. 3. Transition Rates of SIS Model.

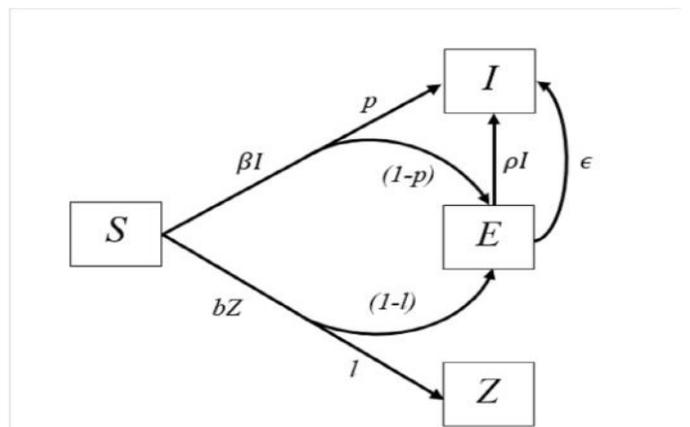


Fig. 4. Transition Rates of SEIZ Model.

TABLE I. NOMENCLATURE FOR STATE TRANSITION VARIABLES OF SEIZ MODEL

Variable	Definition
Susceptible (S)	The onset of the spread of infection
Exposed (E)	Incubator class, in contact with infection
Infected (I)	Adopters class, incubate infection and eventually manifest it and becoming a member of adopter
Skeptic (Z)	susceptible are turned off from infection and become skeptic/stifler
Population(N)	Total population

TABLE II. DESCRIPTION OF PARAMETERS IN SEIZ MODEL

Parameter	Meaning	Units
$\beta$	Rate of contact between S and I	Per unit time
$b$	Rate of contact between S and Z	Per unit time
$\rho$	Rate of contact between E and I	Per unit time
$p$	Transmission rate S->I, given contact with I	Unit-less
$(1-p)$	Transmission rate S->E, given contact with I	Unit-less
$l$	Transmission rate S->Z, given contact with Z	Unit-less
$(1-l)$	Transmission rate S->E, given contact with Z	Unit-less
$\square$	Transmission rate of E->I	Per unit time

In [16], studied the SEIZ and SIS models for the spread of news on Twitter and found the SEIZ model produced much better results. When the models were optimized against Twitter data from 4 news stories and 4 rumors, it was found that the SEIZ model consistently had a much lower relative error than the SIS model. SEIZ suffers from two limitations i.e. the rate of change of population in a state and state transition probabilities considered constant for entire period of observation. This requires the population to reach a certain level before making any trustworthy characterization of information as untrue information resulting in substantial delay in characterization of untrue information. Assuming static rate of population, estimates delayed accuracy of results. In this paper, a dynamic SEIZ computes the rate of change of population at fixed intervals and the predictions based on the new rates periodically. This approach provides the early indications of information being untrue.

### III. MODEL FITTING

The scripts used for fitting the model and performed multi-linear regression to study Twitter data from actual data. The entire experimentation performed in Python 3, solving ODE equation with period 7 days of information spread by Higgs Boson. The computed response ratio for days,  $R_{SI}$  and relative error ( $Err$ ) produced between infected tweets in time,  $I(t)$  and actual tweets. The inputs were taken from Higgs networks which are fitted in each of the compartments of S, E, I, Z, where Z to be skeptic population assumed to be 1 as the initial condition for prevailing days with time.

### A. Dataset used

The Stanford Network Analysis Project (SNAP) used 14 million tweets [33]. With the discovery of a new particle, the Higgs dataset used to analyze the process of information-spread on Twitter. The messages posted on Twitter on this exploration between 1st and 7th July 2012 taken into consideration. Accessibility of the four directive networks derived from Twitter user behaviors is as follows:

- 1) Re-tweeting (retweet network)
- 2) Replying (reply network) to existing tweets
- 3) Mentioning (mention network) other users
- 4) Friends/followers social relationships among user involved in the above activities
- 5) Information about activity on Twitter during the discovery of Higgs boson

It is worth remarking that the user IDs has anonymised, and the same user ID used for all networks. This choice allows using the Higgs dataset in studies about large-scale interdependent/interconnected multiplex/multilayer networks, where one-layer accounts for the social structure and three layers encode different types of user dynamics.

### B. Experimental Evaluation

Data on a particle discovery obtained from Twitter, applied the classic model, and modified model of SEIZ. Here the social network contains 456,631 nodes and 14,855,875 edges to fit and model in dynamic SEIZ. Further its accuracy evaluated for early detection of rumor content by  $R_{SI}$  and  $R_{OIZ}$  (i.e. standard parameters for rumor measure in the network).

In the context of Twitter, various compartment SEIZ interpreted as follows: susceptible (S) is a person who has not heard about the tweet, infected (I) indicates an individual who had tweeted about the tweet, skeptical (Z) are individuals who have learned of the tweet but chooses not to tweet it. In addition, the exposed (E) are those who receive tweets but it took some time before the posting. The inputs for the function are the initial conditions of each compartment(S, E, I, Z), and the total population as  $N = S+E+I+Z$ . Then solving the system of ODEs and computing the result. The initial conditions known of the dataset. Assuming that the skeptic Z (i.e. dormant accounts) to be 1% of total population. The goal of optimizing these parameters is to minimize equation. Some of the constraints to this minimization problem are that all of the parameters from Table III must be non-negative. Thus the parameters values  $\beta, b, \rho, s \geq 0$  and  $0 \leq p, l \leq 1$ . Similarly, the compartments must also be non-negative.

The parameters in Table III calculated from the SEIZ ODE solving previously mentioned in the equations (8-11). To demonstrate SEIZ mathematical model, quantities from Higgs Boson fitted in this model. Data implementation written in Python to extract data from each network from this large social network data set. Here classical SEIZ model computed and compared with the result of modified SEIZ model that provides early detection and indication of rumor in the network provided with information in time slots of 7 days. Experimentation performed for one day, three days, five days, seven days and data obtained for each day respectively.

TABLE III. PARAMETER ESTIMATIONS FOR THE SEIZ MODEL

Duration(Days)	$\beta$	$b$	$\rho$	$s$	$p$	$l$
1 Day	6.1799	4.539E-05	2.1431E-01	0.1356	.9767E-01	5.234E-05
3 Days	.6865	5.0281E-05	.0238	.0452	.3616	19.398E-10
5 Days	2.471	1.8101E-05	.0085	.0271	.0781	4.1900E-10
7 Days	1.260	1.9234E-06	.0044	.0198	.02843	1.526E-12

Fig. 6 demonstrates that the susceptible ( $S$ ), Exposed ( $E$ ), Infected ( $I$ ), Skeptic ( $Z$ ) individuals began to increase with the time (i.e. retweet/reply/mention in the network during the 7 days). For 1 day,  $E(t)$  and  $S(t)$  decreases at the rate negative to that of  $I(t)$  increase. In fact,  $I(t)$  coincides with  $S(t)$  with number of days making more exposed population in  $E(t)$  compartment. For other days, it is observed that  $S(t)$ ,  $I(t)$  and  $E(t)$  becomes constant over certain period in the network. Thus, there is no significant amount of changes seen for 3 days, 5 days and 7 days with static classical SEIZ approach. The remaining cases of SEIZ plots are shown in Fig. 5 with  $S$ ,  $E$ ,  $I$ ,  $Z$  dynamically changing over time.

The time varying network for all four compartments in Higgs Twitter's SEIZ time course plots in Fig. 6 suggest that the effective rate of skeptic population becoming susceptible is much higher than the infected population for 1 day. With an increase in  $Z(t)$ ,  $S(t)$  decreases and thereafter,  $S(t)$  is stable together with  $Z(t)$ .  $I(t)$  increase as  $S(t)$  decreases, but its rate of change is slower. There is also a strong correlation between  $E(t)$  and  $I(t)$  increases. As  $I(t)$  peak,  $E(t)$  peaks as well. However, the increase of  $E(t)$  closely correlated with the decrease compartment  $S(t)$ . Most of the increase  $I(t)$  occurs after  $S(t)$  has developed minimum values, which demonstrates that the infected compartment is constantly changing.

Obtained results from data shows early indication of untrue spread of information/particle spreads which states that dynamic SEIZ works well when obtained in dynamic time intervals over a large network. It also clearly distinguishes data from classical model by stating its rate of change in the entire four networks to be faster with time irrespective of early distinguishing the true and untrue information.

C. Comparison with Classical and Dynamic SEIZ

The modified SEIZ provides early indication of untrue information in the network with increase in number of days in the network. For 5 days to 7 days we can observe significant peak rise with modified technique which suggest that increase in infected users is because of transition of large amount of users from infected to exposed compartment over time ( $t$ ) in the network. Comparing both techniques in Fig. 7.

The infected compartment increases rapidly from the exposed population and not from the direct recruitment of susceptible individuals. Thus, these findings suggest that the generation of early indication of propagation of misinformation spread in the Higgs Boson network.

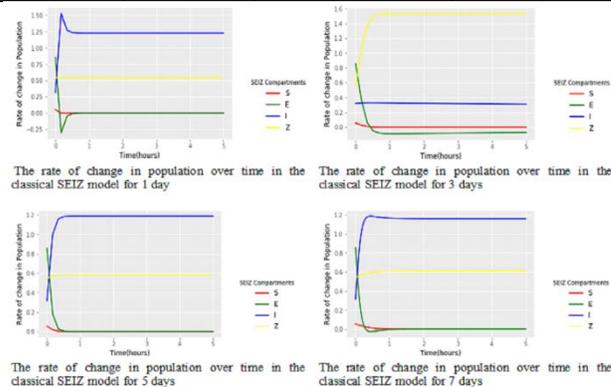


Fig. 5. Rate of Change in Population Over Time for Classical SEIZ Model.

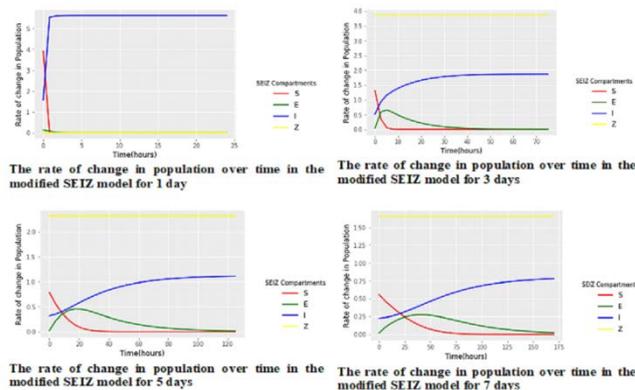


Fig. 6. Rate of Change in Population Over Time in Modified SEIZ Model.

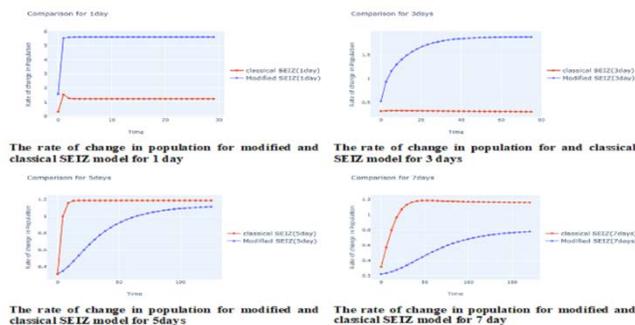


Fig. 7. Comparison between Modified and Classical SEIZ Model Over Time.

#### IV. RUMOR DETECTION

Next requires determining the Exposed Ratio ( $R_{SI}$ ) and the Reproductive Number ( $R_{IZ}$ ) to recognizing the quantity of untrue information on Twitter data. Analysis of this fixed point's local stability shows that the specific reproductive numbers are given by equation 10 where,

$$R_0^{IZ} = \left( \frac{\beta(\varepsilon + p\mu)}{\mu(\varepsilon + p\mu)} \right) \quad (12)$$

$R_0^I$  corresponds to an eigenvector with the adopter component.  $R_0^{IZ}$  corresponds to the exclusive rise of the skeptical population without recognizing the infection.  $R_0^{IZ}$  is the criterion for adopters invaded in a susceptible population. If  $R_0^{IZ} > 1$  it will spread the infection. The higher  $R_0^{IZ}$ , more successful is the acceptance ratio of the infection. By assembling equations to relate the primary parameters of the SEIZ model,  $R_{SI}$  is the ratio of the number of effective transition rates that arrive (from  $S$ ) to the sum of transition levels that exit the compartment (to  $I$ ).  $R_{SI}$  holds all of the SEIZ model's rate constants and probability values and relates them to the exposed flux-ratio i.e. the ratio of effects entering  $E$  to those leaving  $E$ . The ratio, called the Exposed ratio, given by equation 11:

$$R_{SI} = \frac{(1-p)\beta + (1-l)b}{\rho + \varepsilon} \quad (13)$$

The Exposed ratio may be able to indicate a difference between news stories and rumors. A value greater than 1 for RSI means that the transition rates from  $S$  to  $E$  were greater than from  $E$  to  $I$ . If the ratio is less than 1, than the exit rates are greater than the entry rates. For the cases studied in [16], it was shown that the news stories typically had  $RSI > 1$  and the rumors had  $RSI < 1$ . Similar computation followed on dataset with these values indicating early signs of rumor/infection in the network. Table 4 illustrates RSI and RIZ the values for the modified SEIZ model.

The values of RSI in the Fig. 8 shows that for the very first day when SEIZ applied, the exposed population from Susceptible to Infection compartment detects with -15.95 rate of influx to efflux. This shows that modified dynamic SEIZ results in early prediction of rumor count, which initially detected for 1 day. For the rest of days this ratio positively shows with 5 to 6.39 which potentially giving non-rumor rate of diffusion.

$R_0^{IZ}$  corresponds to an eigenvector with the adopter component gives exclusive rise of the skeptical population without recognizing the infection.  $R_0^{IZ}$  is the criterion for adopters invaded in a susceptible population. If  $R_0^{IZ} > 1$  it will spread the infection. The higher  $R_0^{IZ}$ , more successful is the acceptance ratio of the infection. By assembling equations to relate the primary parameters of the SEIZ model,  $R_{SI}$  is the ratio of the number of effective transition rates that arrive (from  $S$ ) to the sum of transition levels that exit the compartment (to  $I$ ).  $R_{SI}$  holds all of the SEIZ model's rate constants and probability values and relates them to the exposed flux-ratio i.e. the ratio of effects entering  $E$  to those

leaving  $E$ . The ratio, called the Exposed ratio, given by equation 13.

Thus,  $R_{SI}$  determines that the greater  $R_{SI}$  value more susceptible enter in exposed group and with the time this ratio decreases which gives us there can be situation that all exposed population can prone to untrue information with time information diffused in the network. For 1 day to 7 days provides a clearer indication of  $I(t)$  infected population creating rumor.

The  $R^{IZ}$  value estimates the average number of adopters induced by a standard spreader in a susceptible population as shown in Fig. 9. These tests the infection's effectiveness as  $R^{IZ} > 1$ , the infection spreads in a network. The greater the  $R^{IZ}$ , the greater is the potency of the infection Fig. 10 determines the similar result that infection rate is much higher for the first day with 122.57 which gradually decreases with  $I(t)$  becoming  $Z(t)$  skeptic. Larger the values of  $R^{IZ}$  estimated in the SEIZ model are mainly due to the very long lifeline of the infection  $I(t)$ . In practice,  $R^{IZ}$  can be estimated by linearizing  $I(t)$  around disease-free equilibrium in simple models [36]. In Fig. 11,  $R^{IZ}$ s provides a statistical distribution to match Feynman diagram spread data [34].

TABLE IV. RUMOR DETECTION RATIO CALCULATED FOR RSI AND RIZ

Time(Days)	$R_{SI}$	$R_{IZ}$
1day	-15.95	122.57
3days	6.35	43
5days	6.39	34
7days	5.05	25

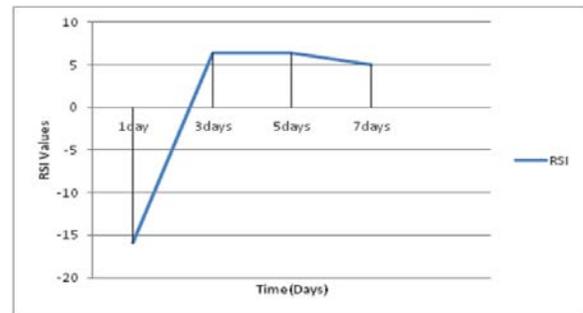


Fig. 8. RSI value for SEIZ.

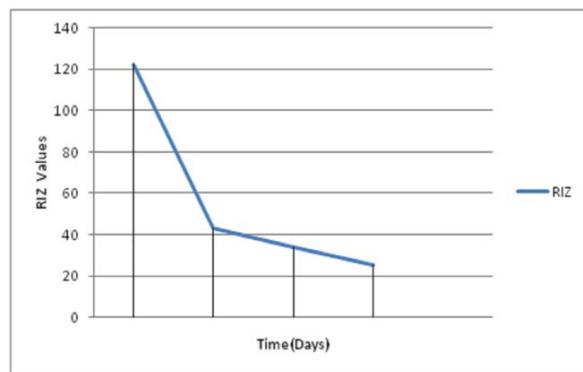


Fig. 9. RIZ value for SEIZ.

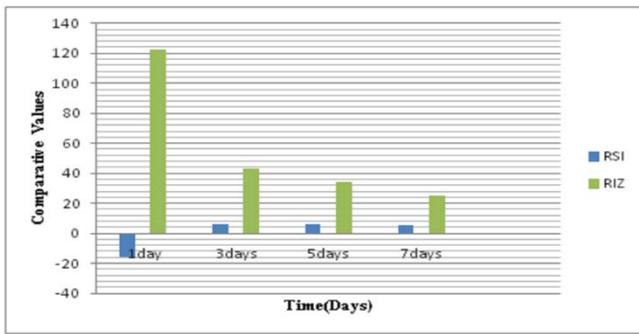


Fig. 10. Comparative value for Rumor Detection.

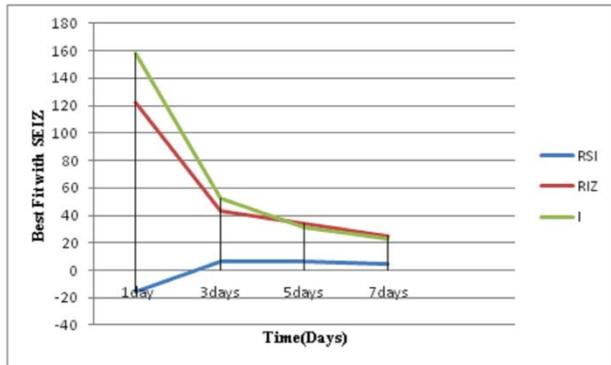


Fig. 11. Best fit in Modified SEIZ.

## V. CONCLUSIONS

Early prediction can remind users of all potential risks with false information before it exists. The false news intervention can eradicate users to erase the negative impacts of false news. In this paper, a model inspired SEIZ epidemiological applied to the population to detect false information dissemination. SEIZ model is capable of modeling the false content of large sets of network with earliest possibility in dynamic network. Exposed ratio ( $R_{SI}$ ) and Reproductive Number ( $R^{IZ}$ ) is a unique value and too sensitive to change in parameter values used to gain useful information from it. Specifically, on Twitter, this paper demonstrates how these criteria incorporated into the strategy to support the early detection of Twitter infection. The findings suggest that Twitter data provides valuable dissemination information with other data analysis strategies to provide accurate and reliable results.

## VI. FUTURE WORK

By examining, more data sets in the future further will strengthen the claim made that SEIZ model is a great model for early prediction of untrue information propagation on Twitter. In addition, these epidemiological models with reinforcement learning can mitigate the effect of false news on social media. Removing suspicious accounts online, protecting users with false information and some automated methods aids in reducing the spread of false news online. Thus has a potential research direction for big data systems.

## REFERENCES

[1] Barnes, J.A., 1954. Class and Committees in a Norwegian Island Parish. *Hum Relations*. <https://doi.org/10.1177/001872675400700102>.

[2] News Industry - Statistics & Facts [WWW Document], n.d. URL <https://www.statista.com/topics/1640/news/>

[3] Social media stats worldwide [WWW Document], n.d. URL <https://gs.statcounter.com/social-media-stats/>

[4] Share of adults who have ever shared fake news or information online in the United States as of March 2019 [WWW Document], n.d. URL <https://www.statista.com/statistics/657111/fake-news-sharing-online/>

[5] Which, if any, is your main source of news? [WWW Document], n.d. URL <https://www.statista.com/statistics/198765/main-source-of-international-news-in-selected-countries/>

[6] Statistics & Facts about Fake News [WWW Document], n.d. URL <https://www.statista.com/topics/3251/fake-news/>

[7] Nguyen, N.P., Yan, G., Thai, M.T., Eidenbenz, S., 2012. Containment of misinformation spread in online social networks, in: Proceedings of the 4th Annual ACM Web Science Conference, WebSci'12. <https://doi.org/10.1145/2380718.2380746>

[8] Guy, M., Earle, P., Ostrum, C., Gruchalla, K., Horvath, S., 2010. Integration and dissemination of citizen reported and seismically derived earthquake information via social network technologies, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). [https://doi.org/10.1007/978-3-642-13062-5\\_6](https://doi.org/10.1007/978-3-642-13062-5_6)

[9] Fan, L., Lu, Z., Wu, W., Thuraisingham, B., Ma, H., Bi, Y., 2013. Least cost rumor blocking in social networks, in: Proceedings - International Conference on Distributed computing Systems. <https://doi.org/10.1109/ICDCS.2013.34>

[10] Budak, C., Agrawal, D., Abbadi, A. El, 2011. Limiting the spread of misinformation in social networks, in: Proceedings of the 20th International Conference on World Wide Web, WWW 2011. <https://doi.org/10.1145/1963405.1963499>

[11] Pastor-Satorras, R., Vespignani, A., 2011. Epidemic spreading in scale-free networks, in: The Structure and Dynamics of Networks. <https://doi.org/10.1515/9781400841356.493>

[12] Newman, M.E.J., 2003. The structure and function of complex networks. *SIAM Rev.* <https://doi.org/10.1137/S003614450342480>

[13] Liu, D.; Yan, E.W. S., 2014. M. Microblog information diffusion: Simulation based on sir model. *J. Beijing Univ. Posts Telecommun.*

[14] Qazvinian, V., Rosengren, E., Radev, D.R., Mei, Q., 2011. Rumor has it: Identifying misinformation in microblogs, in: EMNLP 2011 - Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference.

[15] Bettencourt, L.M.A., Cintrón-Arias, A., Kaiser, D.I., Castillo-Chávez, C., 2006. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Phys. A Stat. Mech. its Appl.* <https://doi.org/10.1016/j.physa.2005.08.083>

[16] Jin, F., Dougherty, E., Saraf, P., Mi, P., Cao, Y., Ramakrishnan, N., 2013. Epidemiological modeling of news and rumors on Twitter, in: Proceedings of the 7th Workshop on Social Network Mining and Analysis, SNA-KDD 2013. <https://doi.org/10.1145/2501025.2501027>

[17] Wang, H., Li, Y., Feng, Z., Feng, L., 2013. Retweeting analysis and prediction in microblogs: an epidemic inspired approach. *China Commun.* <https://doi.org/10.1109/CC.2013.6488827>

[18] Abdullah, S., Wu, X., 2011. An epidemic model for news spreading on twitter, in: Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI. <https://doi.org/10.1109/ICTAI.2011.33>

[19] Gross, T., D'Lima, C.J.D., Blasius, B., 2006. Epidemic dynamics on an adaptive network. *Phys. Rev. Lett.* <https://doi.org/10.1103/PhysRevLett.96.208701>

[20] Wang, Q., Lin, Z., Jin, Y., Cheng, S., Yang, T., 2015. ESIS: Emotion-based spreader-ignorant-stifler model for information diffusion. *Knowledge-Based Syst.* <https://doi.org/10.1016/j.knosys.2015.02.006>

[21] Wang Hao, Li Yiping, Feng Zhuonan, Feng Ling, 2013. ReTweeting Analysis and Prediction in Microblogs: An Epidemic Inspired Approach. *CHINA Commun.*

[22] Stai, E., Milaiou, E., Karyotis, V., Papavassiliou, S., 2018. Temporal Dynamics of Information Diffusion in Twitter: Modeling and Experimentation. *IEEE Trans. Comput. Soc. Syst.* <https://doi.org/10.1109/TCSS.2017.2784184>

- [23] Wang, C., Yang, X.Y., Xu, K., Ma, J.F., 2014. SEIR- based model for the information spreading over SNS. Tien Tzu Hsueh Pao/Acta Electron. Sin. <https://doi.org/10.3969/j.issn.0372-2112.2014.11.031>
- [24] Zhang, Y., Ruan, X., Wang, Haining, Wang, Hui, 2014. What scale of audience a campaign can reach in what price on Twitter?, in: Proceedings - IEEE INFOCOM. <https://doi.org/10.1109/INFOCOM.2014.6848048>
- [25] John, C.; Joshua, S.A., 2014. Epidemiological modeling of online social network dynamics.
- [26] Feng, L., Hu, Y., Li, B., Stanley, H.E., Havlin, S., Braunstein, L.A., 2015. Competing for attention in social media under information overload conditions. PLoS One. <https://doi.org/10.1371/journal.pone.0126090>
- [27] Wang, Y., Zheng, B., 2014. On macro and micro exploration of hashtag diffusion in Twitter, in: ASONAM 2014 - Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. <https://doi.org/10.1109/ASONAM.2014.6921598>
- [28] Daley, D.J., Kendall, D.G., 1964. Epidemics and Rumours [49]. Nature. <https://doi.org/10.1038/2041118a0>
- [29] Castillo, C., Mendoza, M., Poblete, B., 2011. Information credibility on Twitter, in: Proceedings of the 20th International Conference Companion on World Wide Web, WWW 2011. <https://doi.org/10.1145/1963405.1963500>.
- [30] Shah, D., Zaman, T., 2011. Rumors in a network: Who's the culprit? IEEE Trans. Inf. Theory. <https://doi.org/10.1109/TIT.2011.2158885>.
- [31] Isham, V., Harden, S., Nekovee, M., 2010. Stochastic epidemics and rumours on finite random networks. Phys. A Stat. Mech. Its Appl. <https://doi.org/10.1016/j.physa.2009.10.001>.
- [32] Grassly, N.C., Fraser, C., 2008. Mathematical models of infectious disease transmission. Nat. Rev. Microbiol. <https://doi.org/10.1038/nrmicro1845>.
- [33] De Domenico, M., Lima, A., Mougél, P., Musolesi, M., 2013. The anatomy of a scientific rumor. Sci. Rep. <https://doi.org/10.1038/srep02980>.
- [34] Schweber, S.S., Treiman, S., 1994. QED and the Men Who Made It: Dyson, Feynman, Schwinger and Tomonaga Phys. Today. <https://doi.org/10.1063/1.2808749>.
- [35] Estabrook, G.F., 2004. Mathematical Biology: I: An Introduction. Third Edition. Interdisciplinary Applied Mathematics, Volume 17 . By J D Murray. New York: Springer . \$59.95. xxiii + 551 p; ill.; index. ISBN: 0-387- 95223-3. 2002. Mathematical Biology: II: Spatial Models an. Q. Rev. Biol. <https://doi.org/10.1086/421587>.
- [36] Goffman, W., Newill, V.A., 1964. Generalization of epidemic theory: An application to the transmission of ideas. Nature. <https://doi.org/10.1038/204225a0>.