Mathematical modelling and stability analysis of fake news spread with Simulink

PROJECT REPORT

Submitted in partial fulfillment of the requirements
of MATH F491

BITS Pilani
K K Birla Goa Campus

by

Roshan Nair - 2015B4A30344G

Under the supervision of

Dr. Amit Setia

Contents

List of Figures

Acknowledgement

I would like to express out gratitude to Dr.Amit Setia for giving me an opportunity to be a part of this project. The literature survey, mathematical modelling and implementation involved has pushed me to correlate theoretical concepts and get a flavor of the application of essential modelling to other domains. I would like to acknowledge the guidance given by Dr. Setia in routine discussions and his eagerness to help me with my ideas. Lastly, I would like to acknowledge the use of the student license of MATLAB R2018a.

Introduction

Epidemiological Models date back to the late 1920s where Kermack and McKendrick used such models to study the plague. From there, the various generalizations of epidemic models have enabled us to represent most epidemics with a great deal of accuracy. We use the common SEIR epidemic model [1] to model diseases. We essentially hope to use this an inspiration to model news spread as a compartmental model. The rest of this report is organised as follows: We first introduce the model basic SIS model [5] and then we move to the SEIR model [1] and then we discuss the SEIZ fake news model [5]. Simulink implementations are discussed in every case and grpahs as presented as and when necessary.

1 Importance of modelling Fake News Spread

In today's world, with the extreme use of social media the spread of news has become faster then ever. This is certainly the dream, the ability to get to know everything happening in every corner of the world in a few seconds. But, off late the use of these various media to spread misinformation has risen exponentially. There are many examples for this. The very recent WhatsApp scenario where people were sharing forwards with wrong information is one to keep in mind. WhatsApp reacted timely by introducing a restriction on the number of forwards. Whether this actually worked or not is an entirely story altogether.
The necessity to have a model and thereby the ability to visualize the spread of misinformation becomes an important issue in today's day and age. With this ability companies such as WhatsApp itself can find better ways to control the spread of this misinformation. With this is mind we proceed to study the modelling of fake news spread.

2 Use of compartmental models for fake news spread

As mentioned in the introduction the use of compartmental models has been in existence for a very long time. The SEIR model came about for infectious diseases. The section deals with a small explanation as to why these SEIR models can be used for fake news modelling.
The spread of a rumour can be thought of as analogous to the spread of a communicable disease. The population of infected individuals in a disease are similar to the population of individuals who spread a rumour in society. Now, the recovered population is essentially those who no longer believe the rumour is true.

The transfer rate, β is nothing but the rate at which a susceptible individual gets converted to infected ones. Now, given this understanding of why and how fake news is similar to a disease we proceed to some basic fake news models and their stability.

3 Fake News SIS MODEL

$S' = -\beta \frac{SI}{N} + \alpha I$

$I' = \beta \frac{SI}{N} - \alpha I$

Assumptions made in the model and explanation

- A person that actively spreads wrong information about a topic is considered as an infected individual
- A susceptible person is one that is yet to be exposed to this information. A person is said to remain susceptible till he or she comes in contact with an infected individual
- β is the rate at which a susceptible individual gets transferred to an infected individual.
- The population is assumed to be a constant, N.

3.1 Equilibrium Points

The equilibrium points of the SIS model can be found rather easily. We equate the RHS of the S' and I' equations simultaneously and solve for S and I. We get two equilibrium points:

1. Rumour Free Equilibrium = $E_0 = (N, 0)$

2. Endemic Equilibrium= $E_1 = (\frac{\alpha N}{\beta}, N(1-\frac{\alpha}{\beta}))$

The Rumour free equilibrium point is of importance because it gives a way to calculate a quantity called the Basic Reproduction Number.
The Basic Reproduction Number, R_0 is defined as the number of new infected people and existing infected individual gives rise to.
R_0 also helps us analyse the stability of the Rumour free and endemic equilibrium points.

3.2 Reproduction Number

The reproduction number can be calculated from evaluating I'=0 at the rumour free equilibrium point.
We get $\beta\frac{SI}{N} - \alpha I = 0$, which is to be evaluating S at E_0.

This essentially yields the value of I as $\frac{\beta}{\alpha}$, which is nothing but the basic reproduction number.

Therefor,

$$R_0 = \frac{\beta}{\alpha}$$

Now, we will proceed to see in the next section how this R_0 gives rise to conditions which influence the stability of the model at both the equilibrium points.

3.3 Stability Analysis

In order to perform stability analyze, we first have to linearise the system about the equilibrium points. This process is equivalent to finding the Jacobian of the system and evaluating it at the equilibrium points.

3.3.1 Rumour Free Equilibrium

$$J = \begin{pmatrix} -\beta\frac{I}{N} & -\beta\frac{S}{N} + \alpha \\ \beta\frac{I}{N} & \beta\frac{S}{N} - \alpha \end{pmatrix}$$

Now we proceed to evaluate this Jacobian at the rumour free equilibrium. we get the following ;

$$J(E_0) = \begin{pmatrix} 0 & -\beta + \alpha \\ 0 & \beta - \alpha \end{pmatrix}$$

The eigenvalues of this matrix are $\lambda_1 = 0$ and $\lambda_2 = \beta - \alpha$

For a system to be stable at a particular equilibrium point, the eigenvalues that arise from the Jacobian at that point need to be negative. In some cases we relax this condition and say we may have one zero eigenvalue and the rest as negative. Such a system is marginally stable.

We can clearly see that if $\lambda_2 = 0$, the system won't be stable. In order for the system to be stable we need $\lambda_2 < 0$. This happens only when $\frac{\beta}{\alpha} < 1$, which is nothing but the $R_0 < 1$.
Therefore, we see the role that R_0 is playing in the stability of the system.

3.3.2 Endemic equilibrium point

We evaluate the Jacobian calculated in the previous section at the endemic equilibrium and derive the condition for stability.

$$J(E_1) = \begin{pmatrix} -\beta(1-\frac{\alpha}{\beta}) & 0 \\ \beta(1-\frac{\alpha}{\beta}) & 0 \end{pmatrix}$$

The eigenvalues of the Jacobian are :
$\lambda_1 = 0$ and $\lambda_2 = -\beta(1 - \frac{\alpha}{\beta})$

Similar to the previous section we see that for the system to be marginally stable, we must have that $(1 - \frac{\alpha}{\beta}) > 0$, which gives us the condition that $R_0 > 1$
So we can see that based on the value of R_0 only one equilibrium point can be stable at a time.

3.4 Simulink Implementation

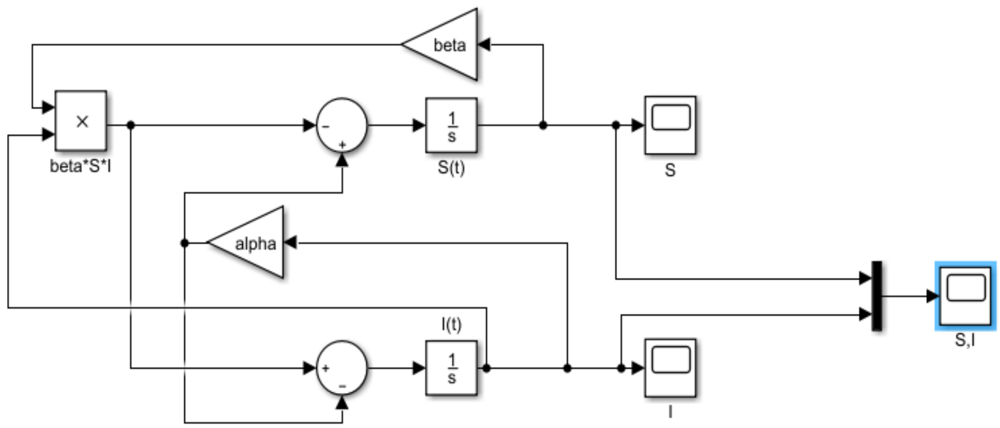

Figure 1: Simulink Model

3.5 Graphs

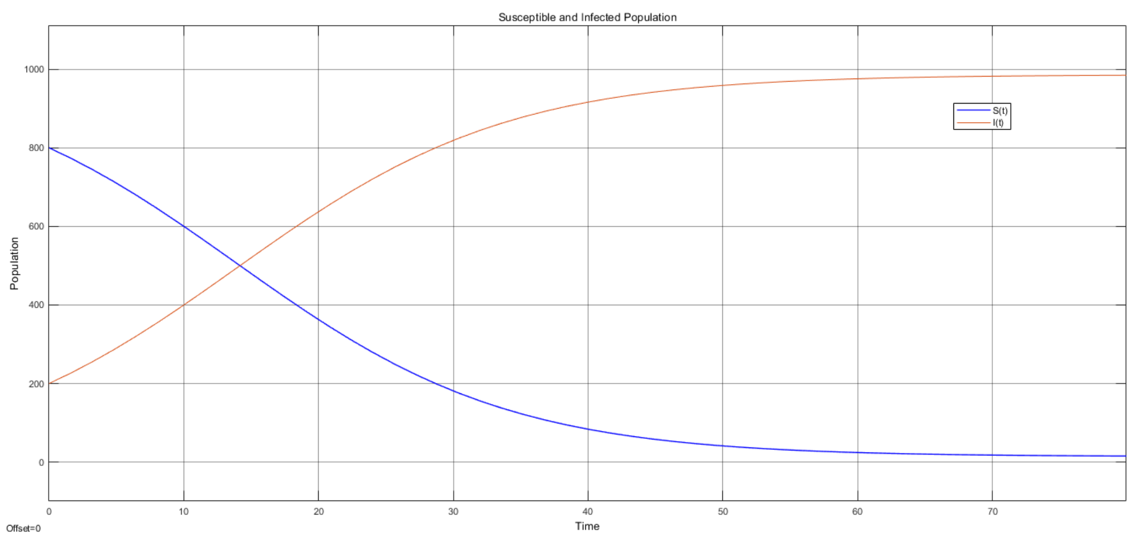

Figure 2: S, I population

4 SEIR model for diseases

4.1 Mathematical Model

The SEIR model with non constant population gives the most general understanding of disease dynamics. S represents the Susceptible population, E represents the Exposed population, I the infected population and R the recovered population. N is the total population at a given time. S,E,I,R and N vary with time and system is best represented by the following equations:

$S' = bN - \beta\frac{SI}{N} - dS + wR$

$E' = \beta\frac{SI}{N} - \sigma E - dE$

$I' = \sigma E - \gamma I - dI$

$R' = \gamma I - dR - wR$

b	Natural Birth Rate
d	Natural Death Rate
w	Rate at which Immunity is lost
β	Transmission Rate from class S to I
γ	Recovery rate
σ	Transmission Rate from class E to I

4.2 Stability Considerations

We study the stability about two equilibrium points that are extremely important for any disease dynamics model. The first is the Disease free equilibrium, where the populations of E, I and R are 0 and $S = \frac{bN}{d}$. The second equilibrium point is termed as the endemic equilibrium point. Both equilibrium points are the solutions we get from setting S', E', I' and R' to 0 simultaneously.

4.2.1 Stability of the Disease Free Equilibrium(DFE)

DFE=$E_0 = (\frac{bN}{d}, 0, 0, 0)$

Now to study the stability of the system at the DFE, we first linearize the system about the DFE. On linearizing the system (Using the Jacobian) and then evaluating the Jacobian at the DFE we in face obtain the A matrix of the corresponding linear system. Proceeding with the standard theory of linear systems, we we can show that this A matrix fact all negative eigenvalues we can conclude that the DFE is a stable equilibrium point.

$$J(S,E,I,R) = \begin{pmatrix} \frac{-\beta I}{N} - d & 0 & \frac{-\beta S}{N} & w \\ \frac{\beta I}{N} & -\sigma - d & \frac{\beta S}{N} & 0 \\ 0 & \sigma & -\gamma - d & 0 \\ 0 & 0 & \gamma & -d - w \end{pmatrix}$$

Now we evaluate this at the DFE for the corresponding linear matrix A.

$$J(DFE) = \begin{pmatrix} -d & 0 & \frac{-\beta b}{d} & w \\ 0 & -\sigma - d & \frac{\beta b}{d} & 0 \\ 0 & \sigma & -\sigma - d & 0 \\ 0 & 0 & \gamma & d - w \end{pmatrix}$$

We further simplify this matrix by taking $\frac{\beta b}{d}$ as l, $\sigma + d$ as f, $\gamma + d$ as g and $d + w$ as h. We now obtain:

$$A = \begin{pmatrix} -d & 0 & -l & w \\ 0 & -f & l & 0 \\ 0 & \sigma & -g & 0 \\ 0 & 0 & \gamma & -h \end{pmatrix}$$

The final step in concluding on the stability of the DFE we have to find the eigenvalues. We do so, by equating the characteristic equation of A.
On doing this we obatin the following eigenvalues:

$\lambda_1 = -d$, $\lambda_2 = -h$, $\lambda_3 = \frac{1}{2} \times (-f - g - \sqrt{f^2 - 2fg + g^2 + 4l\sigma})$ and $\lambda_4 = \frac{1}{2} \times (-f - g + \sqrt{f^2 - 2fg + g^2 + 4l\sigma})$

As all the constants, as mentioned before, are positive values it follows directly that λ_1, λ_2 and λ_3 are indeed negative non-zero quantities. If we can now comment on λ_4 we can conclude on the stability of the DFE.

We take λ_4 and add and subtract $2fg$ inside the square-root. We get $\lambda_4 = \frac{1}{2} \times (-f - g + \sqrt{f^2 + 2fg + g^2 - 4fg + 4l\sigma})$. Simply completing the square inside the square root we can observe that if $-4fg + 4l\sigma$ is less than 0 we in fact have that $\lambda_4 < 0$ thereby guaranteeing stability of the DFE. Now, $-4fg + 4l\sigma < 0$ in actuality is a valid assumption and the ratio we so obtain $\frac{\beta b\sigma}{d(\sigma+d)(\gamma+d)}$ is termed as the basic reproduction number. The basic reproduction number R_0 for a disease free equilibrium point is by definition a positive value lesser than 1.
So we have now proven that the DFE is a stable equilibrium point for the system.

4.2.2 Stability of the Endemic Equilibrium point

To study the stability of the above model we first make a simple transformation for ease of calculation.
We use the transform $sN = S$, $eN = E$, $iN = I$ and $rN = R$. This transformation can be thought off as a conversion from absolute numbers of Susceptible, Exposed, infected and Recovered populations to that of proportions.
Now, it is important to note that on using this transformation we can't simply write $S' = Ns'$. Working with a non-constant population, we must include its effect in this transformation.
Therefore, it is correct to write $S' = s'N + sN' = s'N + (b - d)N$. Similar results follow for E,I,R.

Applying this to the original system, we get

$s' = b - \beta si - bs + wr$

$e' = \beta si - \sigma e - be$

$i' = \sigma e - \gamma i - bi$

$r' = \gamma i - wr - br$

We further simplify this model by clubbing together some parameters. We note that as all the constants are in fact positive values, this clubbing together doesn't make a huge difference for the analysis. We take $f = b + \sigma$, $g = \gamma + b$ and

$h = w + b$ and rewrite the system as follows

$s' = b - \beta si - bs + wr$

$e' = \beta si - fe$

$i' = \sigma e - gi$

$r' = \gamma i - hr$

Similar to the case of studying the stability of the DFE we first find the Jacobian and evaluate the Jacobian at the Endemic Equilibrium point. We first find the equilibrium point by equating s', e', i' and r' to 0 simultaneously.

The Endemic equilibrium point (EE) of the transformed system is
$s^* = \frac{fg}{\beta\sigma}$

$e^* = \frac{-bfg^2h + bgh\beta\sigma}{\beta\sigma(fgh - w\gamma\sigma)}$

$i^* = \frac{-bfgh + bh\beta\sigma}{\beta(fgh - w\gamma\sigma)}$

$r^* = \frac{-bfg\gamma + b\beta\gamma\sigma}{\beta(fgh - w\gamma\sigma)}$

It can clearly be seen that s^* is a positive value. We now proceed to show that e^*, $i*$ and r^* are positive. We do this in order to simplify the process of commenting on the stability.

We first start by looking at the term $(fgh - w\gamma\sigma)$. Expanding f, g and h we have:
$(\sigma + b)(\gamma + b)(w + b) - w\gamma\sigma$
On expanding the terms in the bracket we obtain $(...positiveterms...) + w\gamma\sigma - w\gamma\sigma$ which is clearly positive.
We now look at the numerator in i^*. We have $bh\beta\sigma - bfgh$. We simplify this as follows in order to comment on the sign of the numerator.

$bh \times (\beta\sigma - fg) = (bhfg) \times (\frac{\beta\sigma}{fg} - 1) = (bhfg) \times (R_0\frac{d}{b} - 1)$
Here we note that $R_0 \times p$, where $p < 1$ is the effective reproduction number. In our scenario, $R_0 \times \frac{d}{b}$ is the effective reproduction number, R_{eff}. As mentioned earlier, we need that $R_{eff} > 1$ for the disease to propagate. Under this assumption we have that the numerator of $i^* > 0$. Having already shown the numerator is positive we can conclude that i^* is indeed a positive value.

Using this knowledge we can very easily show that e^* and r^* are also positive values..

Now that we have that all the arguments of the EE are positive we proceed to discuss the stability at the EE. We proceed without substituting the actual values of the EE into the Jacobian. We instead leave the s,e,i,r terms as us but keep in mind that they are indeed positive values. We do this in order to simplify the mathematical expressions we have to deal with.

We also note that as the Jacobian Matrix is a complex at the EE we refrain from explicitly finding the eigenvalues. We instead resort to finding the characteristic equation and using Descartes' Rule of Sign Changes criteria to deduce stability.
We first calculate the Jacobian of the most simplified system obtained in this section.

$$J(s, e, i, r) = \begin{pmatrix} -\beta i & 0 & -\beta s & w \\ \beta i & -f & \beta s & 0 \\ 0 & \sigma & -g & 0 \\ 0 & 0 & \gamma & -h \end{pmatrix}$$

The linearized matrix A at EE is just J(s,e,i,r) evaluated at the EE. We now obtain the characteristic equation for J(EE) as follows:

$$-iw\beta\gamma\sigma + (-h-\lambda) \times [(-g-\lambda) \times (bf + b\lambda + f\lambda + \lambda^2 + fi\beta + i\beta\lambda) - (-bs\beta - s\lambda\beta)\sigma] = 0$$

This characteristic equation of the system linearized about the EE in fact has all negative roots [and we thereby conclude that the system is stable at the EE as well.

4.3 Simulink Implementation with Graphs

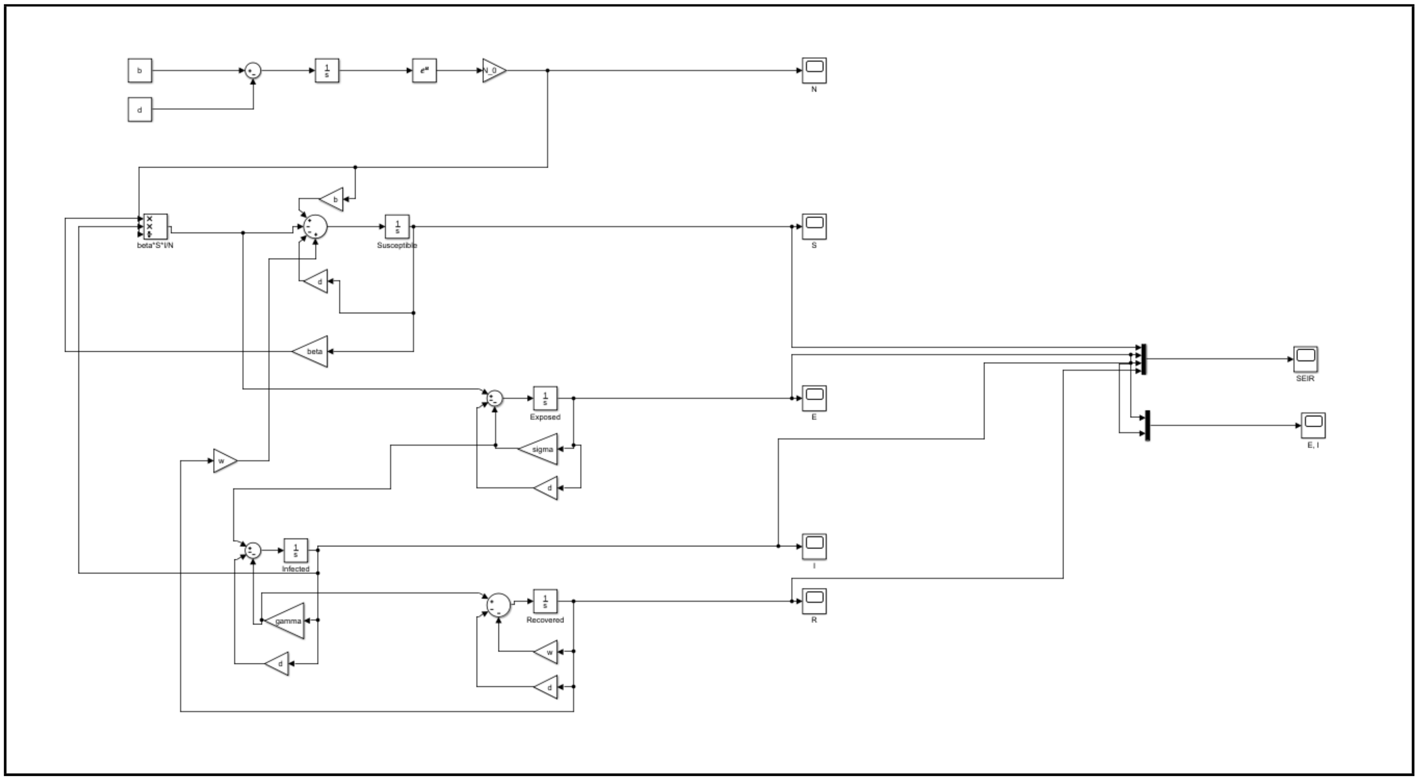

Figure 3: Block Diagram representation of SEIR Model

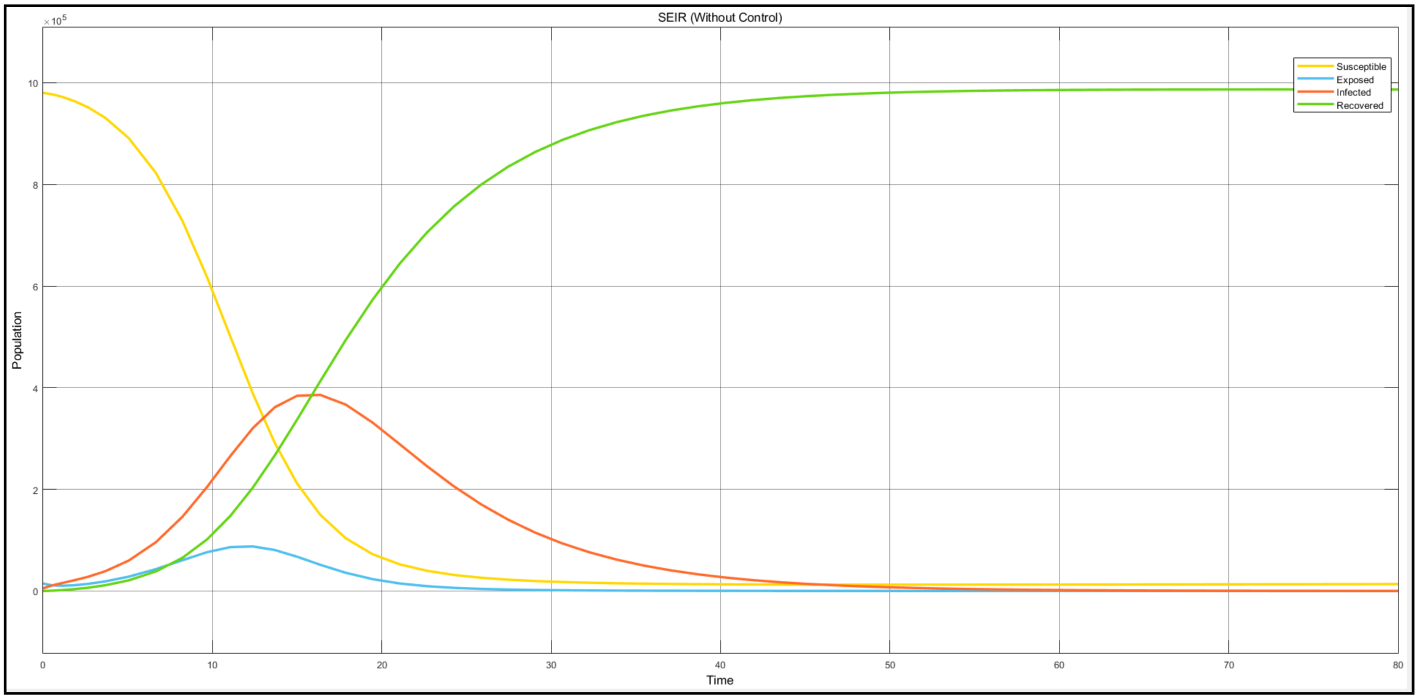

Figure 4: SEIR Dynamics

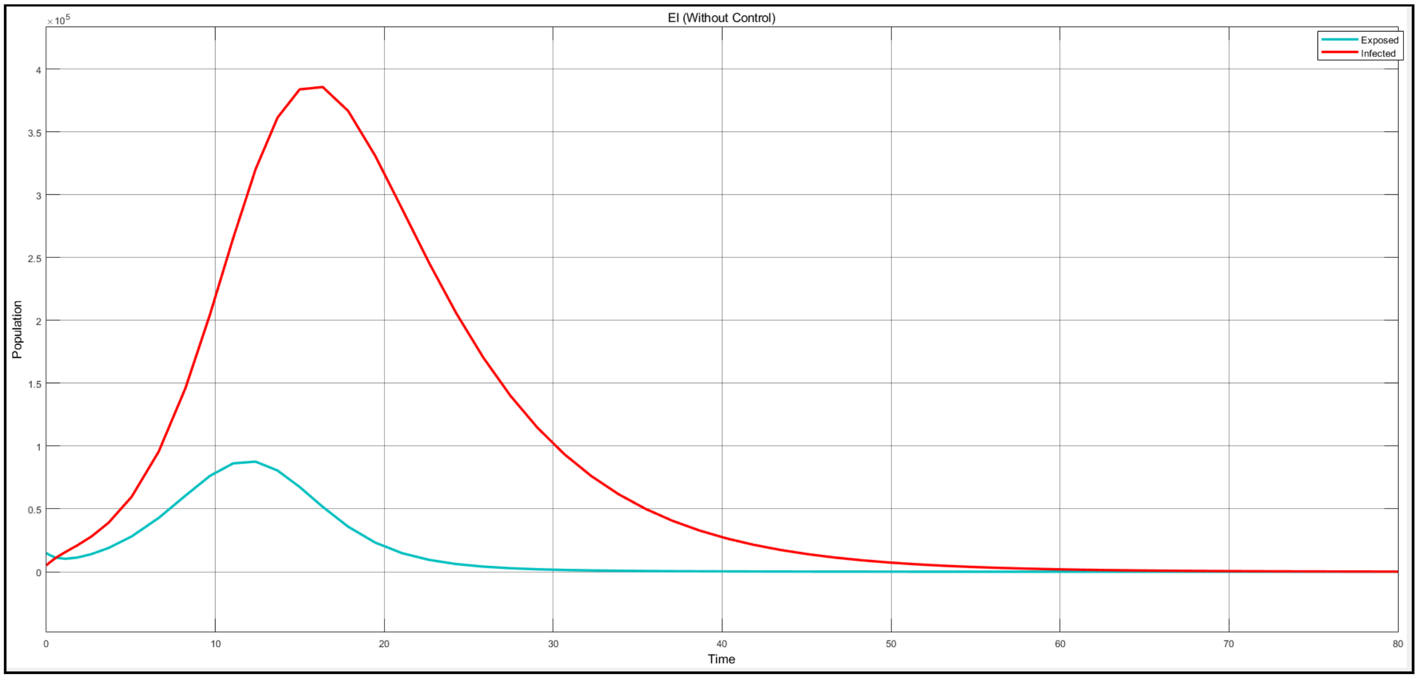

Figure 5: Dynamics of Exposed and Infected Populations

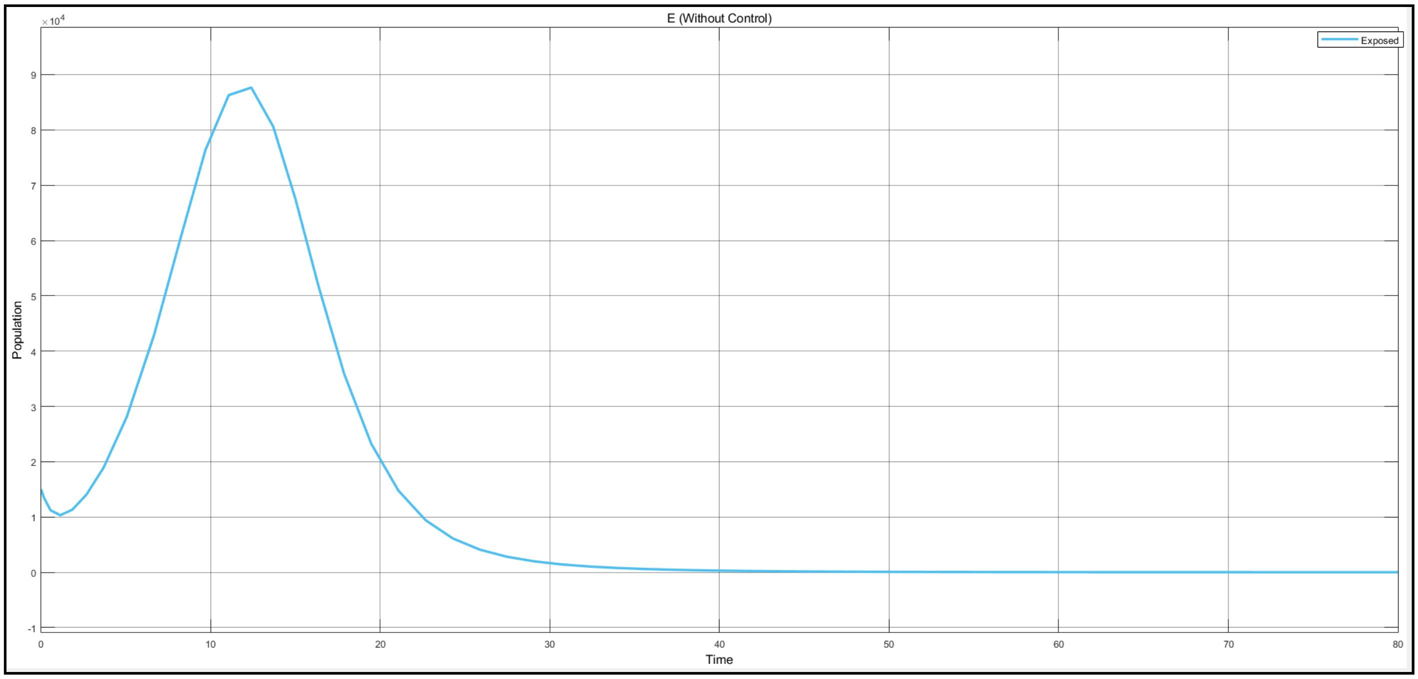

Figure 6: Dynamics of Exposed Population

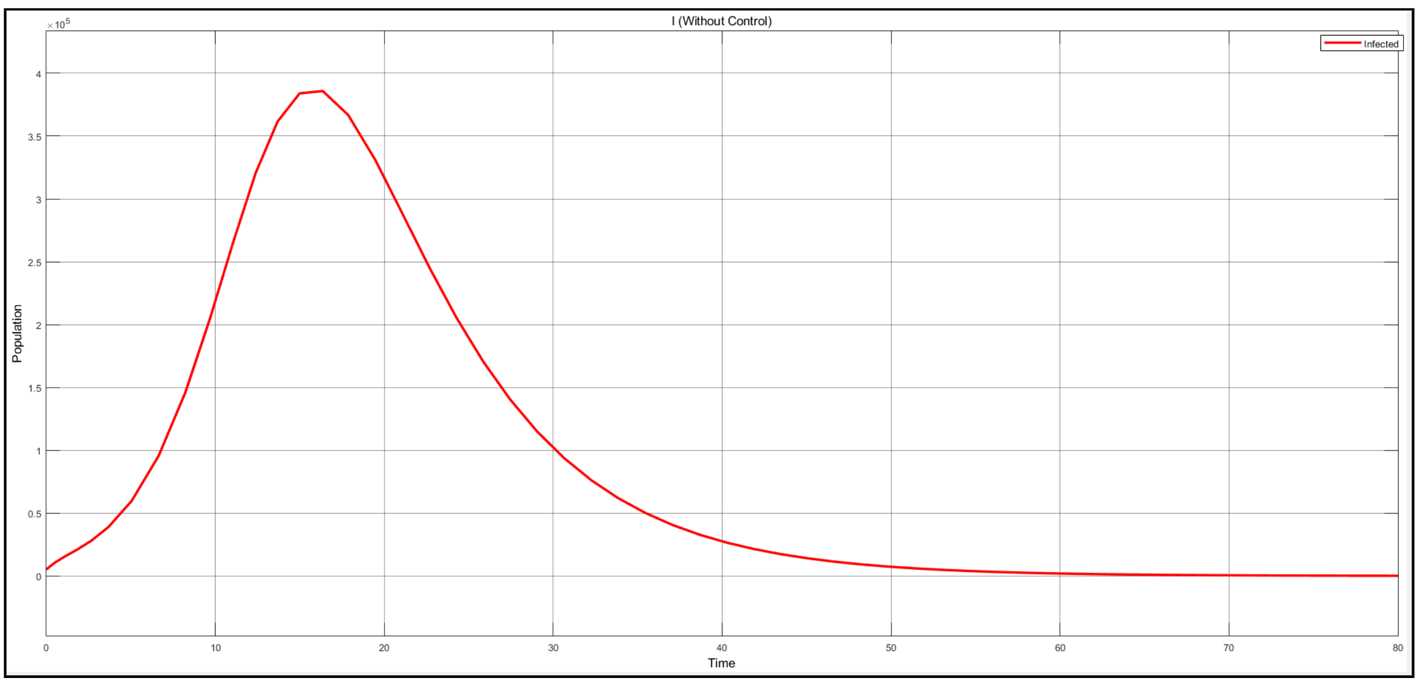

Figure 7: Dynamics of Infected Population

5 Basic SEIZ Model

Having already seen the SIS model we can infer that while that model is a good start towards modelling news spread it lacks detail. We therefore take direct inspiration from the SEIR epidemiological model presented above and go ahead and use it to formulate the SEIZ model.
The compartments in this model are: Susceptible, Exposed, Infected and Z which is essentially those who have realised the news is fake.

$S' = -\beta \frac{SI}{N} - \sigma \frac{SZ}{N}$

$E' = (1-p)\beta \frac{SI}{N} + (1-l)b\frac{SZ}{N} - \rho \frac{EI}{N} - \epsilon E$

$I' = p\beta \frac{SI}{N} + \rho \frac{EI}{N} + \epsilon E$

$Z' = l\sigma \frac{SZ}{N}$

σ	Rate at which Z influences S
β	rate from S to E
ρ	rate at which I influences E
ϵ	rate from E to I
p	proportion of E that turn into I
l	proportion of E that turn to Z

5.1 Stability

In order to discuss the stability of the SEIZ model, we take a simple case assuming $p = 0$ and $l = 1$, the model reduces to :

$S' = -\beta S\frac{I}{N} - \sigma S\frac{Z}{N}$

$E' = \beta S\frac{I}{N} - \rho E\frac{I}{N} - \epsilon E$

$I' = \rho E\frac{I}{N} + \epsilon E$

$Z' = \sigma S\frac{Z}{N}$

Now we discuss the stability about the Rumour free equilibrium point. This is essentially the equilibrium point before which the spread of the rumour starts. That is $S = N$, $E = I = Z = 0$. We can see that as Z is dependent on S, E, I and N as $Z = N - (S + E + I)$ we can neglect it from the state space.

$$J(E_0) = \begin{pmatrix} 0 & 0 & -\beta \\ 0 & -\epsilon & \beta \\ 0 & \epsilon & 0 \end{pmatrix}$$

The eigenvalues of this Jacobian are $\lambda_1 = 0$, $\lambda_2 = \frac{1}{2}(-a - \sqrt{a}\sqrt{a+4b})$ and $\lambda_3 = \frac{1}{2}(-a + \sqrt{a}\sqrt{a+4b})$ where a and b are ϵ and β respectively.
We can clearly see that λ_3 is positive and therefor the system is unstable at the rumour free equilibrium point.

6 Fake News Spread Model with population considerations

The previous model dealt with a constant population and moreover neglected the birth and death of individuals in the system. We now propose an extension of the previous model with this addition. We include the effect of a recruitment rate and the effect of a death rate, both of which we will model as constants for the time being.
The new model is as follows:

$S' = bN - \beta S\frac{I}{N} - \sigma S\frac{Z}{N} - dS$

$E' = \beta S\frac{I}{N} - \rho E\frac{I}{N} - \epsilon E - dE$

$I' = \rho E\frac{I}{N} + \epsilon E - dI$

$Z' = \sigma S\frac{Z}{N} - dZ$

Here b and d represent the birth and death rates (of the population in question) respectively.
Therefore, the population is no longer constant, rather it is given by the following:
$N' = (b - d)N$

This makes the SEIZ model similar to the SEIR model with non-constant population discussed in detail in the previous section. The dynamics and stability considerations are quite similar and are therefore omitted.

7 Future Scope

Having studied the modelling of news spread and made some modifications to existing models,the next logical step is to get data and curve fit these models in order to understand better and more accurately the dynamics of fake news spread.

REFERENCES

[1] Marek B. Trawicki, Deterministic Seirs Epidemic Model for Modeling Vital Dynamics, Vaccinations, and Temporary Immunity, Mathematics 2017, 5, 7; doi:10.3390/math5010007

[2] M. De la Sen, S. Alonso-Quesada, Vaccination strategies based on feedback control techniques for a general SEIR-epidemic model, Applied Mathematics and Computation 218 (2011) 3888–3904

[3] Xuhui Tan, Lingling Yuan, Jingjing Zhou, Yinan Zheng, Fen Yang, Modeling the initial transmission dynamics of influenza A H1N1 in Guangdong Province, China, International Journal of Infectious Diseases 17 (2013) e479-e484

[4] Sarah A. Al-Sheikh, Modeling and Analysis of an SEIR Epidemic Model with a Limited Resource for Treatment, Global Journal of Science Frontier Research Mathematics and Decision Sciences, Volume 12 Issue 14, 2012

[5]Fang Jin, Edward Dougherty , Parang Saraf, Yang Cao, Naren Ramakrishnan, Epidemiological Modeling of News and Rumors on Twitter, The 7th SNA-KDD Workshop '13 (SNA-KDD'13), August 11, 2013, Chicago, United States. Copyright 2013 ACM 978-1-4503-2330-7