Journal of Computing Research and Innovation (JCRINN) Vol. 6 No. 2 (2021) (pp55-65)
https://jcrinn.com : eISSN: 2600-8793
doi: 10.24191/jcrinn.v6i2.200

Susceptible-Infected-Recovered (SIR) Model to Measure the Virality of Breaking News on Facebook

Noorzila Sharif[1], Jasmani Bidin[2*], Ku Azlina Ku Akil[3], Nur Natasha Arisha Rizan[4]
[1,2,3,4] Faculty of Computer and Mathematical Sciences,
Universiti Teknologi MARA Perlis Branch, Arau Campus, 02600 Arau, Malaysia

Corresponding author: jasmani@uitm.edu.my*
Received Date: 28 January 2021
Accepted Date: 5 February 2021
Revised Date: 12 March 2021
Published Date: 1 April 2021

HIGHLIGHTS

- SIR models are not only reliable in measuring the spread of any contagious diseases but also reliable in measuring the spread of news in social media.
- Based on the value of initial reproduction number , $R_0 > 1$, it shows that the breaking news is viral.
- Based on curves of the graphs, the dissemination of selected news related to Covid-19 content has increased and reached the peak within 4 days. It means that the news spread faster at the beginning of the incident has been posted in Facebook (outbreak).
- This preliminary study could be extended and correlated to many other factors such as stock market responses, national productivity and employment rate. Further studies on different types news contents such as politics, economy, business and entertainment should also be considered because the degree of disseminations may be different. The analysis will be more meaningful if different transmission parameter values are tested to identify an expectation after a certain action has been taken.

ABSTRACT

Susceptible-Infected-Recovered (SIR) model has been used worldwide to measure the spreading of covid-19 in the community. Apparently, the spreading nature of the covid-19 virus and any other contagious disease is quite similar with the spreading of breaking news through social media. This study was carried out to analyze the dynamics spread of one selected news content on Facebook using SIR models with demography and without demography. From the news, the numbers of likes, comments, shares, views as well as the number of followers of the Facebook account have been collected to calculate reproduction number. For SIR without demography, the reproduction number (R_o) is 1.69, indicates that for every 100 Facebook users who received the news, they will probably share the news to other 169 Facebook users. The value of R_0 is slightly lower (1.58) for SIR with demography. This preliminary study could be extended by considering a lot more observations and by testing different parameters value due to any further action imposed after the news spreading out.

***Keywords**: breaking news, demography, Facebook, SIR model, reproduction number*

Journal of Computing Research and Innovation (JCRINN) Vol. 6 No. 2 (2021) (pp55-65)
https://jcrinn.com : eISSN: 2600-8793
doi: 10.24191/jcrinn.v6i2.200

INTRODUCTION

Social media expands our reach more quickly, much further, and at a grander scale through words, pictures, and videos. Social media networking sites such as Twitter, Facebook, YouTube, and Instagram empower individuals to share their voice in a media-centric. These form of social media allow the users not only receive any news from anybody having the social account but also allow the users to disseminate the news to others very rapidly (Al-Rawi, 2019). Social media makes interaction easier and much quicker than real-life interaction. With the flying success of the internet and the creation of social media, social networking symbolizes the podium in which people interact, collaborate and communicate (Khurana & Kumar, 2018). During this global attack of pandemic coronavirus (Covid-19), social media plays an important role in the society to disseminate the related news although some of them are fake news. Social media has the ability in which the key information is used appropriately and properly to provide fast and efficient distribution routes (Chan et al., 2020).

Facebook has become the most popular social network worldwide (Statista, 2019). It has become a primary source of news sharing (Baek, Holton, Harp & Yaschur, 2011). More than 30 billion pieces of content are uploaded each month, such as web links and news stories (Glynn, Huge, & Hoffman, 2012). Network structure in Facebook encourages content to be shared and discussed. The sharing process is simple. Moreover, Facebook also allows users to have a group discussion. Other than that, status updates are a new form of one-to-many interactions that users of social networking sites frequently use (Hum et al., 2011). Facebook users can update the status by sharing the thoughts on a particular issue. By doing so, this will indirectly share the information to other Facebook users.

Facebook acts as an information medium since it provides the opportunities in distributing the news on sports, health, politics, entertainment, natural disaster and many more. The breaking news, the top current story or event, on Facebook normally has an increase in the number of likes, comments, shares and views day by day. This breaking new is being highlighted and Facebook users will share the news and repost the news on their Facebook page. The news is widely disseminated and it spreads very fast. However, the dynamic of the news spreading and the duration of the news to hit the maximum numbers of viewers are uncertain and need to be studied.

This study attempted to analyze the dynamics involved in the dissemination of the breaking news content on Facebook. The mathematical solution method offered in this study involved formulating a model of the spreading of breaking news on Facebook and to compare the growth and the decline of the number of viewers in relation to breaking news based on Susceptible-Infected-Recovered (SIR) models with demography and without demography.

SIR model is originally used to investigate the spread of contagious disease such as dengue (Asmaidi et al. 2014; Side and Salmi, 2013) and tuberculosis (Side et al., 2017; Kalu and Inyama, 2012). The appearance of Covid-19, a new rapidly widespread contagious disease in early 2020, has urged a lot of researchers to use the SIR model to investigate its spreading nature (Cooper et al. 2020; Alanazi et al., 2020). In addition to the disease outbreak, the application of SIR models has also been extended to different fields (Ji, Lu, Yeung & Hu, 2017; Zaman & Jung, 2007). Since the news spreading nature is similar to disease endemic, this study used the Susceptible-Infected-Recovered (SIR) model to investigate news spread via Facebook.

Journal of Computing Research and Innovation (JCRINN) Vol. 6 No. 2 (2021) (pp55-65)
https://jcrinn.com : eISSN: 2600-8793
doi: 10.24191/jcrinn.v6i2.200

METHODOLOGY

For this preliminary study, the breaking news on Facebook related to Covid-19 in China, has been selected. The news is about China expanding its lockdown because the virus has spread around the world. The news has been observed after 5 days it is released on Facebook account from Cable News Network (CNN). The information about the number of users disclosed to this viral content, the number of Facebook users receiving and sharing the content and the number of Facebook users stop posting the content have been observed from 29th January 2020 until 12th February 2020 (14 days). The data are obtained from the number of likes, shares, comments and views shown in the content of the news. The data will be analyzed using SIR model without demographic and with demographic. The SIR model with demography includes the birth or death rate where the birth rate is assumed to be equal to the death rate (Rodrigues, 2016). Meanwhile, SIR model without demography is the model that does not include the death or birth rates.

i- SIR Model without Demography

In SIR model without demography, the Facebook population is divided into three classes which are susceptible, infected and recovered as shown in Figure 1.

Figure 1: SIR Model without demography

Each of the compartment is presented as:

S – Susceptible (Facebook user disclosed to viral content),

I – Infected (Facebook user receiving and sharing the viral content),

R – Recovered (Facebook user stops posting the viral content).

Parameter β is the rate of transmission of the number of Facebook users who receive and share the viral content among those who are disclosed to the content. Then, γ is the recovery rate of the number of Facebook users stop posting among those who are receiving and sharing the content. Thus, the compartments change over time, *t,*

S (t) is the number of Facebook user disclosed to viral content at time, *t*,

I (t) is the number of Facebook user receiving and sharing the viral content at time, *t* and

R (t) is the number of Facebook user stops posting the viral content at time, *t*.

Journal of Computing Research and Innovation (JCRINN) Vol. 6 No. 2 (2021) (pp55-65)
https://jcrinn.com : eISSN: 2600-8793
doi: 10.24191/jcrinn.v6i2.200

The total population, N at time t is

$$N(t) = S(t) + I(t) + R(t). \quad (1)$$

It is assumed that all Facebook users are active users, no new user creates the account and no user deletes the Facebook account during this period. This study also assumed that the rate of decrease of $S(t)$ is proportional to the product of Facebook user disclosed to viral content and Facebook user receiving and sharing the viral content, that is,

$$\frac{dS}{dt} = -\beta S(t)I(t). \quad (2)$$

This study also assumed that the rate of change of Facebook user stops posting the viral content is proportional to the Facebook user receiving and sharing the viral content, that is,

$$\frac{dR}{dt} = \gamma I(t). \quad (3)$$

From Eqn. (1)

$$\frac{dI}{dt} = -\frac{dS}{dt} - \frac{dR}{dt}. \quad (4)$$

Substituting Eqns. (2) and (3) in (4), will lead to

$$\frac{dI}{dt} = \beta S(t)I(t) - \gamma I(t) = I(t)(\beta S(t) - \gamma). \quad (5)$$

Eqns. (2) – (5) will result the followings:

$$\frac{dS}{dt} = -\beta SI, \quad (6)$$

$$\frac{dI}{dt} = \beta SI - \gamma I, \quad (7)$$

$$\frac{dR}{dt} = \gamma I. \quad (8)$$

The transmission rate, β is the average number of Facebook users who have not viewed the content to the number of Facebook users who have already viewed the content such that,

$$\beta = \left(1 - \frac{\text{number of viewers}}{\text{Facebook users}}\right) \times \left(\frac{\text{summation of likes, comments and shares}}{\text{Number of viewers}}\right). \quad (9)$$

Journal of Computing Research and Innovation (JCRINN) Vol. 6 No. 2 (2021) (pp55-65)
https://jcrinn.com : eISSN: 2600-8793
doi: 10.24191/jcrinn.v6i2.200

Meanwhile, the recovery rate, γ is related to the expected duration of the viral content complete (D),

$$\gamma = \frac{1}{D}. \tag{10}$$

Therefore, the formula for reproduction number without demography is

$$R_0 = \frac{\beta}{\gamma}. \tag{11}$$

ii- SIR Model with Demography

In SIR model with demography, the population to be considered is the same as the previous model; susceptible, infected and recovered. The expression of SIR model with demography is also very similar to expression of SIR without demography, except that the inflow and outflow of birth or death rate, μ are added as in Figure 2.

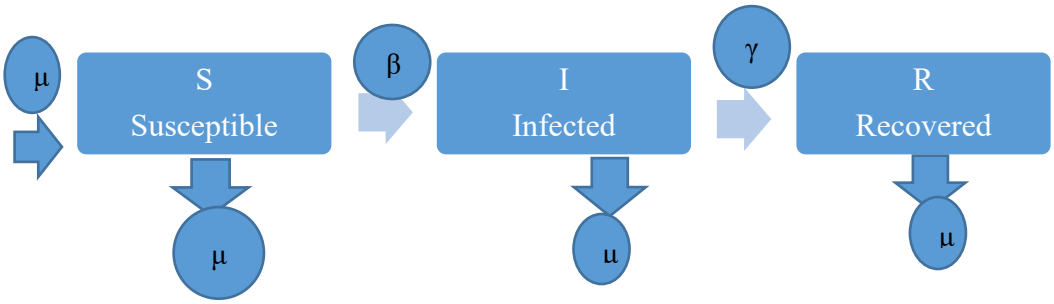

Figure 2: SIR Model with demography

The equations of the SIR model with demography are given as

$$\frac{dS}{dt} = \mu - \mu S - \beta SI. \tag{12}$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I. \tag{13}$$

$$\frac{dR}{dt} = \gamma I - \mu R. \tag{14}$$

where μ is the death rate of population which is equal to the birth rate.

Journal of Computing Research and Innovation (JCRINN) Vol. 6 No. 2 (2021) (pp55-65)
https://jcrinn.com : eISSN: 2600-8793
doi: 10.24191/jcrinn.v6i2.200

Parameter β is the average of number Facebook user who does not view the content yet to the number Facebook user who already view the content.

$$\beta = \left(1 - \frac{\text{number of viewers}}{\text{Facebook users}}\right) \times \left(\frac{\text{summation of likes, comments and shares}}{\text{Number of viewers}}\right). \tag{15}$$

For parameter γ,

$$\gamma = \frac{1}{D}, \tag{16}$$

where D is the duration of the viral content is expected to be completed.

The reproduction basic number indicated by R_0 is explained as the number of average transmissions associated with β (Facebook user receiving and sharing the viral content and the Facebook user number being disclosed to viral content) and $\gamma + \mu$ (Facebook user receiving and sharing the viral content that stops posting the viral content). The formula for reproduction number with demography is

$$R_o = \frac{\beta}{\gamma + \mu}. \tag{17}$$

If $R_o < 1$, the Facebook user number receiving as well as sharing the viral content decreases.

If $R_0 > 1$, the Facebook user number receiving as well as sharing the viral content increases.

If $R_0 = 1$, the Facebook user number receiving as well as sharing the viral content is constant.

Equilibrium Points of the Model

These two equations below are explained at the same time if R does not affect the S and I.

$$\mu - \beta SI - \mu S = 0. \tag{18}$$

$$\beta SI - \gamma I - \mu I = 0. \tag{19}$$

From Eqn. (19),

$$\text{I} = 0 \text{ and } S = \frac{\mu + \gamma}{\beta} = \frac{1}{R_0}. \tag{20}$$

Substitute I and S into Eq. (18) to create equilibrium points that are

$$(S,I)=(1,0) \text{ and} \tag{21}$$

$$(S, I) = \left(\frac{1}{R_0}, \frac{\mu(R_0 - 1)}{\beta}\right). \tag{22}$$

Journal of Computing Research and Innovation (JCRINN) Vol. 6 No. 2 (2021) (pp55-65)
https://jcrinn.com : eISSN: 2600-8793
doi: 10.24191/jcrinn.v6i2.200

The equilibrium points $(S, I) = (1, 0)$ is the viral-free equilibrium since $I = 0$ and the point $(S, I) = \left(\frac{1}{R_0}, \frac{\mu(R_0 - 1)}{\beta}\right)$ is endemic equilibrium since $I \neq 0$.

The data is run by using Maple software in analyzing the SIR model's qualitative behaviour.

FINDINGS AND DISCUSSIONS

The number of likes, comments, shares and views in the posted page, as well as the number of followers of the Facebook account for the Cable News Network (CNN) for the selected breaking news related to Coronavirus (Covid-19) were 19800, 2400, 21800, 358000 and 33100000 respectively as shown in Table 1. The data was collected after 5 days the news were posted.

Table 1: The number of Likes, Comments, Shares, Viewers and Followers of Facebook Account in Malaysia for the selected News

Variables	Coronavirus (Covid19) CNN
Number of likes	19800
Number of comments	2400
Number of Shares	21800
Number of viewers	358000
Number of Followers	33100000

Parameter β is the rate of transmission of a Facebook users receiving and sharing the viral content after being disclosed to viral content. Parameter γ is the rate of recovery of a Facebook users when they stop posting the viral content after receiving and sharing the content and μ is refer to the rate of change the number of Facebook accounts based on the birth or death rate of the population in Malaysia for the year 2018.

The parameters used for both SIR models without demography and with demography are summarized in Table 2.

Journal of Computing Research and Innovation (JCRINN) Vol. 6 No. 2 (2021) (pp55-65)
https://jcrinn.com : eISSN: 2600-8793
doi: 10.24191/jcrinn.v6i2.200

Table 2: Parameter value of SIR Model for Selected News

Parameter	Coronavirus (Covid-19)
β	0.1209
γ	0.0714
μ	0.0053

The number of daily active Facebook users in Malaysia is considered as total population, N. Hence, the total population (Facebook users), N is 23.1 (in million). The initial conditions for the news related to Covid-19 in China are:

$S(0) = 23.06,$

$I(0) = 0.04,$

$R(0) = 0.$

Table 3 represents the basic reproduction number, R_0 for SIR model without demography and SIR model with demography for the selected news. It represents the number of secondary Facebook users receiving the same news. In this analysis, the R_0 for both models are greater than 1 ($R_0 > 1$). It indicates that the news is outbreak and disseminated to other Facebook users. The spread of news is slightly higher through the Facebook user for SIR model without demography as compared to with demography because the death rate is not considered. The $R_0 = 1.69$ indicates that for every 100 Facebook users receiving the news, they will disseminate the news to other 169 Facebook users. From another model, $R_0=1.58$ indicates that for every 100 Facebook users receiving the news, they will share the news to the other 158 Facebook users.

Table 3: The Reproduction Number, R_0

Model	R_0
SIR without demography	1.6933
SIR with demography	1.5762

Figure 3 and Figure 4 show that the Facebook user receiving and sharing the viral content (infected) of covid-19 reaches the peak within four days for both models. It means that the news spread faster at the beginning of the incident happen (outbreak). Facebook users were very distressed and immediately share the news content to others.

Journal of Computing Research and Innovation (JCRINN) Vol. 6 No. 2 (2021) (pp55-65)
https://jcrinn.com : eISSN: 2600-8793
doi: 10.24191/jcrinn.v6i2.200

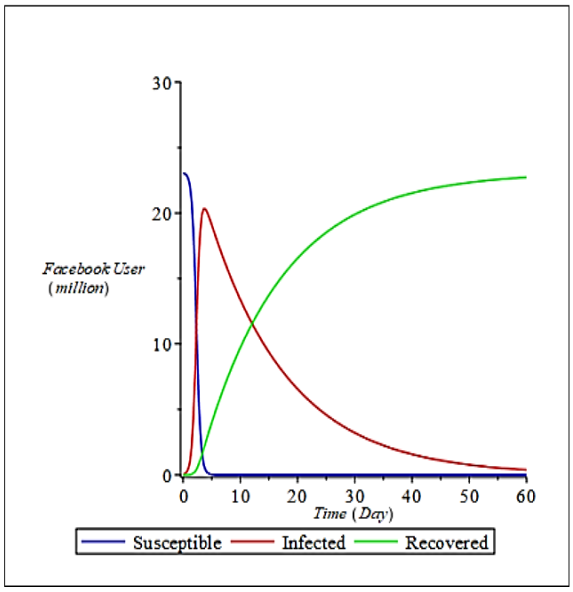

Figure 3: SIR Model without Demography

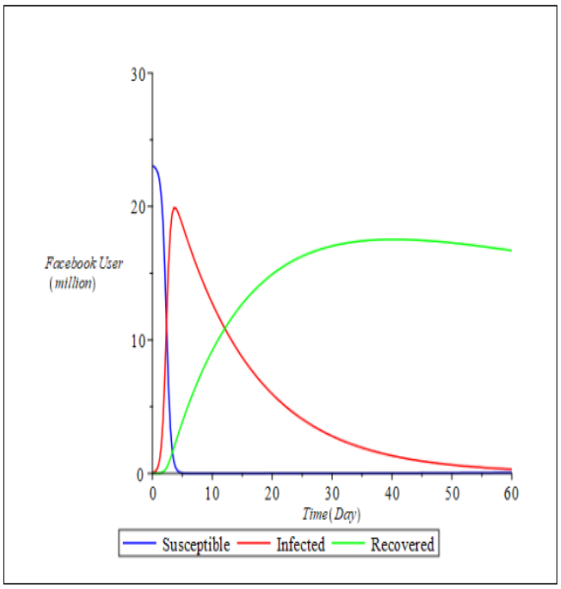

Figure 4: SIR Model Demography

Based on the graph in Figure 3, the number of Facebook users who exposed to the news during the first day of observation is about 22.06 million. On that time, the number of users who received the news is still low (0.04 million). For the next four days, the number of users who received and shared the news increased tremendously until it reached the maximum number (20.4 million). By that time, the news has reached almost all of the users (89% of the users). After day four, the number of users receiving and sharing started to decrease gradually until it reaches 0.02 million on the day 60. In the same time of day 60, the number of users who are no longer interested to the news has reached the maximum numbers of 22.8 million.

The graph in Figure 4 demonstrates that not much difference can be seen for the first 4 days as compared to graph in Figure 3. The only difference can be seen in the number of Facebook users who stopped sharing the news after 60 days of observation (16.8 million). Thus, the number of Facebook users who are no longer interested to the news is less than the number in the previous model (about 6 million). The intersection between Infected and Recovery (day 14) in both figures show that the number of users who received and shared the news decreases sharply, while the number of Facebook users who stopped sharing the news increases sharply. At this time, there is no more Facebook users disclosed to the viral news.

The selected breaking news is said to be epidemics since the infected curve from both graphs exhibits an exponential growth behavior in the early stages of an outbreak. The breaking news about the spread of covid-19 in China had distressed and persuaded Facebooks users to share the news immediately after receiving it since any news related to Covid-19 in China on that time always become prime focus. After day 4, Facebook users get used to the news and start to pay attention to another latest news. As a result, the dissemination rate is decreasing until the number of Facebook users sharing the news is getting very low on day 60. Meanwhile, the recovering rate keeps on increasing very fast before it starts to slowing down on day 30.

CONCLUSION AND RECOMMENDATIONS

The SIR models seems very suitable in analyzing the dynamics of the spreading of breaking news content since it shares the same features or attributes as seen in the dissemination of contagious disease. In this preliminary study, it is found that the selected news related to Covid-19 in China is viral and epidemic. The reproduction number is greater than 1 ($R_o > 1$) indicates that the selected news is outbreak and kept spreading. The spread was faster on the early phase of the incident. It took only 4 days to hit the maximum numbers of Facebook users receiving the news.

The magnitude of reproduction number, R_o is a key in determining the spread of the news. Since R_0 is calculated using a simple SIR model, it has limitation in explaining the detail of the outbreak of the news. In this study, the calculation of R_o is based on the observation in the fifth day after the news has been released. For the future study, it is more appropriate to make the observation right after the news is posted to get more accurate prediction. This analysis uses day to represent time unit. Since the significance of time units is not well known in the context of social networks, identification of the best unit of time to be used might be explored. In the same time, the real data should also be considered to justify the reliability of the model.

For future research, it is also recommended to extend the study to different form of breaking news on Facebook in order to know which one get the most attention from the Facebook users. This study also could be extended and correlated to many factors such as stock market responses, national productivity and employment rate.

ACKNOWLEDGMENTS

The authors express their gratitude to the colleagues and research team members for their collaboration and discussions, which have improved methodologies presented in this work.

CONFLICT OF INTERESTS DECLARATION

The authors declare no conflict of interests regarding the publication of this article.

REFERENCES

Al-Rawi, A. (2019). Viral news on social media. *Digital journalism*, *7*(1), 63-79.

Asmaidi, Siantura, P., & Nugrahani, E. H. (2014). A SIR mathematical model of dengue transmission and its simulation. *IQSR Journal of Mathematics,* 10(2)*,* 56-65.

Alanazi, S. A., Kamruzzaman, M. M., Alruwaili, M., Alshammari, N., Alqahtani, S. A., & Karime, A. (2020). Measuring and preventing COVID-19 using the SIR model and machine learning in Smart Health Care. *Journal of Healthcare Engineering*, Vol. 2020, Article ID 8857346. https://doi.org/10.1155/2020/8857346

Baek, K., Holton, A., Harp, D., & Yaschur, C. (2011). The links that bind: Uncovering novel motivations for linking on Facebook. *Computers in Human Behavior*, *27*(6), 2243-2248.

Chan, A. K. M., Nickson, C. P., Rudolph, J. W., Lee, A., & Joynt, G. M. (2020). Social media for rapid knowledge dissemination: Early experience from the COVID-19 pandemic. *Anaesthesia*.

Cooper, I., Mondal, A., & Antonopoulos,C. G. (2020). A SIR model assumption for the spread of COVID-19 in different communities. *Chaos, solitons, and fractals*, 139, 110057. https://doi.org/10.1016/j.chaos.2020.110057.

Glynn, C. J., Huge, M. E., & Hoffman, L. H. (2012). All the news that's fit to post: A profile of news use on social networking sites. *Computers in Human Behavior*, *28*(1), 113-119.

Hum, N. J., Chamberlin, P. E., Hambright, B. L., Portwood, A. C., Schat, A. C., & Bevan, J. L. (2011). A picture is worth a thousand words: A content analysis of Facebook profile photographs. Computers in Human Behavior, 27(5), 1828-1833.

Ji, S., Lü, L., Yeung, C. H., & Hu, Y. (2017). Effective spreading from multiple leaders identified by percolation in the susceptible-infected-recovered (SIR) model. *New Journal of Physics*, *19*(7), 073020.

Kalu, A. U., & Inyama, S.C. Mathematical model of the role of vaccination and treatment on the transmission dynamics of tuberculosis. *Generation Mathematics Notes*, *11*, 10-23.

Khurana, P., & Kumar, D. (2018, April). Sir Model for Fake News Spreading Through Whatsapp. In *Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIoTCT)* (pp. 26-27).

Rodrigues, H. S. (2016). Application of SIR epidemiological model: new trends. *arXiv preprint arXiv:1611.02565*.

Side, S., & Salmi, M. N. (2013). A SIR model for spread of dengue fever disease (Simulation for South Sulawesi, Indonesia and Selangor, Malaysia). *World Journal of Modelling and Simulation*, *9*(*2*), 96-105.

Side, S., Sanusi, W., Aidid, M. K., & Sidjara, S. (2017). Global stability of SIR and SIER model for tuberculosis disease transmission with Lyapunov function method. *Asian Journal of Applied Sciences*, *9*(*3*), 87-96.

Statista (2019). Statistics. https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

Zaman, G. & Jung, I.H. (2007, December). Stability techniques in SIR epidemic models. In *Proceedings in Applied Mathematics and Mechanics*, Vol 7, No 1, pp 20300632030064. Berlin: WILEY-VCH Verlag.