Susceptible Exposed Infected Resisted (SEIR) Models On Hoax Spread In Twitter

Andreas Rony Wijaya[1,a], Yosef Ronaldo Lete Boro[1,b]
[1a,b] Mathematics Department, Faculty Mathematics and Natural Science
Universitas Sebelas Maret
Surakarta, Indonesia
andreasronywijaya@gmail.com

Harum Sekar Prativi
Mathematics Department, Faculty Mathematics and Natural Science
Medan State University
Medan, Indonesia
harumsekarprativi@gmail.com

***Abstract*— In the digital era and technological development, the internet and social media have become a main thing for the development of information diffusion processes, but at the same time are also used for hoax spread, misinformation, even hoax news. Twitter is a social media that is gaining a high level of popularity. This clearly shows that the spread of news and hoaxes on Twitter takes place very quickly and massively. This hoax can cause disputes between Twitter users. Therefore a policy is needed to reduce the spread of hoaxes. The hoax countermeasure policy can be done appropriately if policy makers know the dynamics of hoax spread. The mathematical model in the form of a system of differential equations can be used to observe the spread of hoaxes on Twitter. In this paper, the author uses the susceptible exposed infected resisted (SEIR) model to analogize the hoax spread model on Twitter. The purpose of this article is to analogize SEIR on the spread of hoaxes on Twitter and determine the solution of the model. This model classifies the population into four individual groups, namely tweets (susceptible), groups of tweets that contain hoaxes and need time (exposed), groups of tweets containing hoaxes and spread (infected), and groups of tweets that are free from tweets (resisted) The SEIR model in the form of a nonlinear differential equations which is difficult to determine the exact completion, therefore it is used numerically, namely the fourth order Runge-Kutta method. From the discussion obtained two equilibrium points are disease-free equilibrium point and endemic equilibrium point. The hoax spread on twitter shows that the equilibrium point was reached on 399 days.**

Keywords-component; SEIR model; the point of equilibrium; Twitter

I. INTRODUCTION

The development of technology and information is currently very rapid in human life. Technology provides convenience for humans in various fields of life and provides many benefits. However, if technology is not used properly, technology can have a negative impact on life. On the one hand, humans can master the world with technology, but on the other hand technology can also dominate humans.

One of the technologies that humans use to exchange information is social media. Social media today seems to be a primary need for humans. But this social media will have a negative impact if it is not used properly. One of the abuses of social media is the spread of hoax information or mere hoaxes on social media. This hoax information can cause confusion between social media users. Zhang (2009) in his article shows that the rapid rate of information transmission can be utilized in an emergency such as a warning of a disaster. Meanwhile, on the negative side, it can lead to destructive hoaxes and can cause social panic and even economic turmoil in a country.

One of the most commonly used social media is Twitter. In the first quarter of 2017, statistics showed that active Twitter users each month reached 317 million users with the number of tweets generated per day reaching around 500 million tweets. This shows that hoaxes will spread faster through social media. Because of the impact of a large hoax, a policy is needed to reduce the spread of hoaxes. The hoax countermeasure policy can be done appropriately if policy makers know the dynamics of hoax distribution. Problems regarding the dissemination of hoax information can be represented and resolved with the Mathematical model. The mathematical model in the form of a system of differential equations can be used to observe the spread of hoaxes on Twitter.

The study of the Mathematical model which discusses the mechanism of the spread of hoax began in 1964 with the classic epidemic model of SIR introduced by Daley and Kendal (1964). In this model, the population is divided into three groups: individuals who know the hoax and spread it (spreader / infected), individuals who do not know the hoax but are vulnerable to spread it (ignorant / susceptible), and individuals who know the hoax but choose not to spread it (stifler / recovery). Then, this DK epidemic model was further developed by other researchers such as Maki and Thompson (1973) who perfected the DK model into an MT model by adding the assumption that if spreader / infected individuals make contact with other individual spreaders then the individual can turn into an individual stifler / recovery.

ISBN 978-602-53524-0-9

Kermack and Mc Kendrick introduced susceptible exposed infected recovery in the spread of several diseases. The results show that the SEIR model can be used to model the spread of ideas with a long process of adoption of ideas requires a long period of time. Based on this description, in this paper, the SEIR model is used to model the spread of hoaxes on twitter and to determine the equilibrium point of SEIR model hoax spread on twitter.

II. LITERATURE REVIEW

A. *Related Research*

Mathematical modeling for the spread of the disease not only provides important information about the spread of the disease into the human chain, but also offer insight into the strategies that can be used to control it. Classification of the human population into different groups formed the basic premise of the use of epidemiological models for modeling the diffusion of information. Two widely used model is a model of SIR (Susceptible, Infected, Recovered) and SIS (Susceptible, Infected, Susceptible). Newman et al. (1996) shows that the model of SIR, can be used for a wider network of variations and complex. Kimura et al. (2002) propose the application of SIS model to study the diffusion of information in which the nodes can be crossed several times.

Liu et al. (2016) proposed a model of the propagation of hoaxes SIHR (Spreaders, Ignorants, Hibernators, Removed), by forgetting and remembering mechanism to simulate propagation of hoaxes in the network that are not homogeneous. Xiong et al. (2007) proposed a model of diffusion with four different circumstances: susceptible, contacted, infected, and refractory (SCIR) to identify the value of the threshold rate of almost zero approaching deployment. Bettencourt et al. (2006) proposed a model of SEIZ (susceptible, exposed, infected, skeptic) to catch Feynman diagram adoption by using a number of publication after World War II. They extract the common feature to spread ideas and predict the adoption process ideas. Their results showed that the model SEIZ can adjust the adoption process with the long-term idea is errors that make sense, but does not indicate whether this model can be applied to large scale datasets, or whether it can be applied on Twitter, where the story unfolds in real time. This will be noticed on the research pattern of spreading hoax on Twitter and will be used to figure out the pattern of the SEIR model its spread.

B. *System of Differential Equation*

According to Boyce and DiPrima on, nonlinear systems of three order differential equation one can generally be rendered as

$$\frac{dx_1}{dt} = f_1(t, x_1, x_2, x_3)$$

$$\frac{dx_2}{dt} = f_2(t, x_1, x_2, x_3)$$

$$\frac{dx_3}{dt} = f_3(t, x_1, x_2, x_3) \qquad (2.1)$$

With the variables x_1, x_2, bound and free variables are x_3 and t. Exact resolution of the system (2.1) is not always easily determined. If its exact settlement difficult, then it can be determined the settlement its approach numerically. Completion of the numerical approach can be used if the system (1) meet the initial value. If the system of equations (1) given the value of the initial $(t_0; a_1; a_2; a_3)$, then the system of equations (2.1) can be expressed as

$$\frac{dx_1}{dt} = f_1(t, x_1, x_2, x_3), \qquad x_1(t_0) = a_1$$

$$\frac{dx_2}{dt} = f_2(t, x_1, x_2, x_3), \qquad x_2(t_0) = a_2 \qquad (2.2)$$

$$\frac{dx_3}{dt} = f_3(t, x_1, x_2, x_3), \qquad x_3(t_0) = a_3$$

Systems of equations (2.2) referred to as the initial value problem. Runge-Kutta method algorithm order four can be used to determine the completion of the approach of an initial value problem.

C. *Susceptibel Infected Recovered (SIR) Model*

The population in this analogy are divided into three groups, namely the groups of individual Tweets (susceptible), group tweets containing a hoax and has already spread (infected), the Group of the non-hoax tweet (resisted). Model of epidemic disease of diabetes and its complications can be written as

$$\frac{dS}{dt} = \mu N - \beta S \frac{1}{N} - \mu S$$

$$\frac{dI}{dt} = \beta S \frac{1}{N} - \gamma I - \mu I$$

$$\frac{dR}{dt} = \gamma I - \mu R \qquad (2.3)$$

More $S(0) > 0, E(0) > 0, I(0) > 0, R(0) > 0$ and μ, β, γ are parameters that is positive. (2.3) model is the one order differential equations system.

ISBN 978-602-53524-0-9

III. METHODS

Here are eight steps which will be taken in the analysis of the data :

1) Identify a pattern of propagation of hoaxes on Twitter.
2) Specify the assumptions, add the variable E, and adds the parameter model SIR on the system (2.3).
3) Determine the relationship (the instantaneous rate of change) between variables based on assumptions and parameters obtained in step (2).

Step (1) - (3) done to reach the goal first.

4) Determine the resolution of the model that have been obtained in step (3), the completion of the model States the dispersion pattern "hoax on Twitter.

Step (4) is carried out to achieve the second objective.

5) Determine the value of each parameter based on data (secondary).
6) Determine the resolution of the model based on the parameter values obtained in step (5) by Runge-Kutta method order four.
7) To draw the graph of the solution obtained in step (6).
8) Interpret graphs completion obtained in step (7).

Step (5) - (8) done for the third goal.

IV. RESULT AND DISCUSSION

A. *SEIR Model*

Spread hoaxes on twitter may be analogous with the model SEIR. The population in this analogy is divided into four individual groups, namely the Group of Tweets (susceptible), group tweets containing a hoax and it took time to spread out (exposed), group tweets containing a hoax and has already spread (infected), group Tweet the non-hoax (resisted). The assumptions used in the analogy model SEIR was :

1. The large number of Tweets a constant.
2. The rate of occurrence and the loss of the same tweet.
3. There's only one hoax that is transmitted on the model populasitweet.
4. Which is exposed by the hoax (E) spreader population can be turned into a hoax (I) and can be turned into a population that resisted against hoax (R).

The large number of Tweets at time t for each group can be expressed as $S(t), E(t), I(t)$ and $R(t)$. Given the assumption of constant population numbers with $N = S(t) + E(t) + I(t) + R(t)$, so that the rate of occurrence and the loss of the same tweet. If it is the rate of occurrence and the rate of disappearance of the tweet, then the large number of occurrences of a tweet of N and the number of missing tweets for each individual group of S, E, I, and R.

The S Group of individuals was reduced from the large number of individual early μN, because there is an infected individual hoax and entered into the individual exposed of $\beta S \frac{1}{N}$ individuals into groups resisted μS, and loss of μS.

$$\frac{dS}{dt} = \mu N - \beta S \frac{1}{N} - \gamma S - \mu S \qquad (4.1)$$

Group E is derived from individual group S of $\beta S \frac{1}{N}$ and reduced because there are individuals who can spread of σE as well as loss of individuals in the Group E resisted μE.

$$\frac{dE}{dt} = \beta S \frac{1}{N} - \sigma E - \mu E \qquad (4.2)$$

Furthermore, the individual on group I come from individuals who entered from the Group E resisted σE. The number of individuals in the Group I is reduced because there is entering into the individual R of λI and the individual lost of μI.

$$\frac{dI}{dt} = \sigma E - \lambda I - \mu I \qquad (4.3)$$

Individual group of R is obtained from the individual S of γS and individuals I of λI, as well as reduced as there are individuals who lost big as μR.

$$\frac{dR}{dt} = \gamma S - \mu R + \lambda I \qquad (4.4)$$

The transfer of a group of individuals can be seen in Figure 4.1.

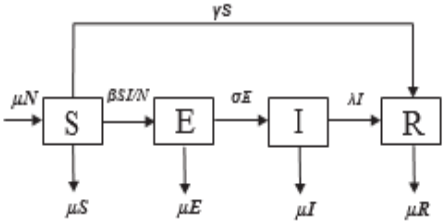

Figure 4.1. Diagram of transfer model SEIR on the spread hoaxes on twitter.

ISBN 978-602-53524-0-9

Thus, from equations (4.1), (4.2), (4.3) and (4.4) retrieved model SEIR on the spread hoaxes on twitter that is

$$\frac{dS}{dt} = \mu N - \beta S\frac{1}{N} - \gamma S - \mu S$$

$$\frac{dE}{dt} = \beta S\frac{1}{N} - \sigma E - \mu E$$

$$\frac{dI}{dt} = \sigma E - \lambda I - \mu I$$

$$\frac{dR}{dt} = \gamma S - \mu R + \lambda I \qquad (4.5)$$

More $S(0) > 0, E(0) > 0, I(0) > 0, R(0) > 0$ and $\mu, \beta, \gamma, \sigma, \lambda,$ are parameters that is positive. $\mu, \beta, \gamma, \sigma, \lambda,$ successively declared lajukemunculan/loss of tweet, tweet a displacement rate of contacts into the contact rate hoax, E with I, the rate of contact I contact rate with R, S with R. .. The pattern of spread hoaxes on twitter can be determined from the settlement model (4.5) which is a system of nonlinear equations.

B. The Equilibrium Point of SEIR Model

Haberman (2005) explains that a point is said to be in equilibrium if the State system does not suffer changes all the time. Equilibrium is a State where a large number of individual changes on the Group S, E, I, and R all the time is zero. The point of equilibrium models (4.5) can be achieved when

$$\mu N - \beta S\frac{1}{N} - \gamma S - \mu S = 0$$

$$\beta S\frac{1}{N} - \sigma E - \mu E = 0$$

$$\sigma E - \lambda I - \mu I = 0$$

$$\gamma S - \mu R + \lambda I = 0 \qquad (4.6)$$

Different settlement system (4.6) produces two types of equilibrium point.

1) Non equilibrium point hoax

$$X_0 = (S_0, E_0, I_0, R_0) = \left(\frac{N\mu}{\gamma+\mu}, 0, 0, \frac{N\gamma}{\gamma+\mu}\right)$$

With $E_0 = 0$ and $L_0 = 0$, it indicates that no individual in the group exposed and infected that spread the hoax.

2) Endemic equilibrium point

$$X_e = (S_e, E_e, I_e, R_e)$$

With

$$S_e = \frac{N(\lambda+\mu)(\mu+\sigma)}{\beta\sigma}$$

$$E_e = -\frac{N\left(\gamma(\lambda+\mu)(\mu+\sigma) + \mu\left(\mu^2 - \beta\sigma + \mu\sigma + \lambda(\mu+\sigma)\right)\right)}{\beta\sigma(\mu+\sigma)}$$

$$I_e = -\frac{N\left(\gamma(\lambda+\mu)(\mu+\sigma) + \mu\left(\mu^2 - \beta\sigma + \mu\sigma + \lambda(\mu+\sigma)\right)\right)}{\beta(\lambda+\mu)(\mu+\sigma)}$$

$$R_e = \frac{N\left(\gamma(\lambda+\mu)(\mu+\sigma)(\lambda+\mu+\sigma) - \lambda\sigma\left(\mu^2 - \beta\sigma + \mu\sigma + \lambda(\mu+\sigma)\right)\right)}{\beta(\lambda+\mu)\sigma(\mu+\sigma)}$$

The value of E_e and I_e are not zero indicates that there are still individuals exposed and infected affecd by the hoax.

C. Application to the Spread Hoax on Twitter

Model SEIR applied to spread hoaxes on twitter. Twitter is a social networking service that allows its users to send and read text-based messages known as Tweets. Tweet is what we write or post in twitter. Tweet can spread through a retweet and mention, i.e. when the username create a tweet and tweet retweet another user's.

The application of model SEIR on the spread hoaxes on twitter on initial conditions there are 85375503 individuals susceptible individual,7322132 exposed, 6518822 individuals infected individual,783543 resisted and has a value of N as much as 800 million. Given the rate of 0.10840, the rate of β and σ of γ, the rate of 0.07423 of 0.01874 of 0.05556, the rate μ, and λ of 0.23152. parameter values that have manuals by model (4.5). Thus, the retrieved model SEIR on the spread hoaxes on twitter that is

$$\frac{dS}{dt} = 0.05556N - 0.10840\frac{SI}{N} - 0.01874S - 0.05556S$$

$$\frac{dE}{dt} = 0.10840S\frac{I}{N} - 0.07423E - 0.05556E$$

$$\frac{dI}{dt} = 0.07423E - 0.23152I - 0.05556I$$

$$\frac{dR}{dt} = 0.01874S - 0.05556R + 0.23152I \qquad (4.7)$$

Models (4.7) is a system of nonlinear differential equations of order one the exact settlement difficult. Therefore, the determined completion of approaches with the Runge-Kutta method requiring four order initial value. The initial value that is used is the number of individual groups of S, E, I, R on the first day of November 6, 2017, i.e.

ISBN 978-602-53524-0-9

$$(S_0, E_0, I_0, R_0) = (85375503, 7322132, 6518822, 783543) \quad (4.8)$$

Furthermore, the specified resolution of the model (4.7) and initial values (4.8). The settlement pattern of the spread hoaxes on twitter. Spread pattern is obtained is presented in figure 4.2.

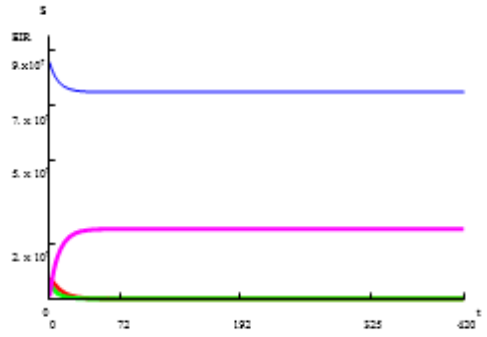

Figure 4.2. The large number of individual S (blue), E (red), I (green), and R (purple)

Based on Figure 2 can see that the Group S experienced a decline, with the decline that is 7.47779 × 10^7 pada day 140 and does not change on the following day. Then the Group E experienced a decline, with the decline of IE 0 day-399 and unchanged today. Furthermore, the Group I experienced a decline, with the decline of IE 0 day-389 and does not change on the following day. In contrast to the R groups has increased with the increase that is 2.52221 × 10^7 pada day 136 does not change on the following day.

According to Haberman (2005), the point of equilibrium is said to be stable if the given initial values that are pretty close to the point of equilibrium is obtained values fairly close to the equilibrium point. The stability of the equilibrium point can be determined denganmengamati the trajectory of the field phase on the model. The point of equilibrium (E, I) = (0, 0). Field phase to the point of equilibrium (E, I) can be seen in Figure 4.3.

From Figure 4.3 seen that point of equilibrium (E, I) is the point that is stable, so the system can be interpreted as the start of the 399 S, E, I, and R respectively stable at (7.47779 × 10^7, 0, 0, 2.52221 × 10^7).

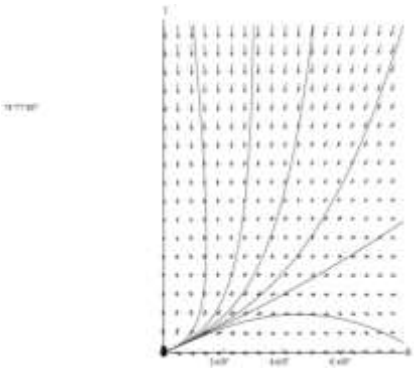

Figure 4.3. Field phase E and I

V. CONCLUSIONS

1. Model SEIR on the spread hoaxes on twitter is written as

$$\frac{dS}{dt} = \mu N - \frac{\beta SI}{N} - \gamma S - \mu S$$
$$\frac{dE}{dt} = \frac{\beta SI}{N} - \sigma E - \mu E$$
$$\frac{dI}{dt} = \sigma E - \lambda I - \mu I$$
$$\frac{dR}{dt} = \gamma S + \lambda I - \mu R$$

More $S(0) > 0, E(0) > 0, I(0) > 0, R(0) > 0$ and $\mu, \beta, \gamma, \sigma, \lambda,$ are parameters that is positive. $\mu, \beta, \gamma, \sigma, \lambda,$ and successive appearance/disappearance rate stated tweet, tweet a displacement rate of contacts into the contact rate hoax, E with I, the rate of contact I contact rate with R, S with R.

2. There are two types of equilibrium point model SEIR on the spread hoaxes on twitter, that is the point of equilibrium is the equilibrium point and hoaxes are non endemic.
3. The application of model SEIR on the spread hoaxes on twitter pointed out that the point of equilibrium is reached on day-399 (7:47779 x 107, 0, 0, 2.52221 x 107). By observing the trajectory of the field phase of the model, it brings the stability point of equilibrium (E; I) = (0; 0).

REFERENCES

[1] Bettencourt, L., 2006, *The Power of A Good Idea: Quantitative Modeling of The Spread of Ideas from Epidemiological Models*, Physica A.
[2] Daley, D. J. &Kendall, D. G., 1904, *Epidemics and Rumours, Nature.*
[3] Haberman, R., 2005, *Mathematical Models, Society for Industrial and Applied Mathematics.*

ISBN 978-602-53524-0-9

[4] Liu, Y., Su-Meng D., & Yi-Xiang Z. 2016.*SHIR competitive information diffusion model for online social media.*

[5] Maki, D. &Thompson, M., 1973, *Mathematical Models and Applications*, Prentice-Hall.Journal Physica A. 461.pp:543-553

[6] Thomas, S. A., 2007, Lies, Damn *Lies, and Hoaxs: An Analysis of Collective Efficacy, Hoaxs, and Fear in The Wake of Katrina*, Sociol Spectrum

[7] Zhang, Z. L., &Zhang Z. Q., 2009, *An Interplay Model for Hoax Spreading and Emergency Development*, Physica

ISBN 978-602-53524-0-9