International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.3 (2016), pp.125-132
http://dx.doi.org/10.14257/ijmue.2016.11.3.12

Dynamical Behavior of Rumor in Online Social Networks

[1]Hui Xie, [1]Yongjie Yan and [2]Yingchun Hou

[1]School of Mathematics & Computer Science, Jiangxi Science & Technology Normal University, Nanchang 330038, P. R. China
[2]Department of Computer Technology, Shangqiu Polytechnic, Shangqiu 476000, P. R. China

Abstract

As a powerful tool for phase transition analysis, the mean-field theory of rumor spreading brings us a new method for research in Online Social Networks (OSN). Based on the theory of complex networks, we simulate the complex process of rumor spreading in OSN. Using numerical simulations with focus on the information spreading threshold as well as critical behaviors, we study the Susceptible-Infected-Recovered- Susceptible (SIRS) information dynamics model in OSN with asymmetric propagation. Basic principle and structure of OSN were analyzed. We present a rumor spreading model, where the influence of the neighbor trees is treated in a more realistic way and the definition of a neighborhood can be tuned by an additional parameter. Based on the mean-field theory approach, a contact network model with scale-free property is built. It is demonstrated that the asymmetry of propagation plays important role: we could redistribute the asymmetry to balance the degree heterogeneity of the network and then to restore the information threshold to a finite value. The relationship of spreading probability, network size, infected fraction and infected ratio are showed respectively in the paper. Our model exhibits a surprisingly sharp phase transition which can be shifted by a redefinition of the neighborhood. The result shows that the SIRS model built in this paper is valid and the simulation of the information propagation is feasible.

***Keywords**: phase transition; online social networks; mean-field theory; epidemic threshold*

1. Introduction

Many networks are conjectured to have a scale-free topology, including the Internet, biological networks, world wide web links, peer-to-peer networks, power grids, protein networks and Online Social Networks (OSN). Recently, much attention has been attracted by OSN from many disciplines in computer science, physics and sociology. In OSN, an identity is a set of descriptive attributes instead of a single identity string. Barabasi and Albert developed scale-free networks to mimic dynamical mechanism which present in real-world networks [1-2]. There is a great significance for using scale-free networks model to analysis of OSN [3]. Analysis of OSN is an important method to help us understanding the nature of interactions between individuals in a society. Studies on OSN provide a useful framework for analyzing social events such as failure analysis [3], opinion dynamics [4] and information spreading [5-9]. OSN dynamics is routinely described as a static network with dynamical aspects, or temporal changes, occurring on the network [10].

In [11], the authors use stochastic process to simulate worm propagation in computer networks, and find that the information threshold for the spread of viruses in scale-free networks is very close to zero. In [12], the authors introduce an e-epidemic SEIR (susceptible-Exposed-Infectious-Recovered) model for the transmission of worms in a computer network is developed to have a better understanding of the reason for cyber war

A virus propagation model of computer, HSIRS propagation model, which is suitable for the cloud environment is proposed in [13]. The stability of the worms-free and endemic equilibrium of the epidemic models is well discussed in [14], and it is also shown that the models may undergo a forward bifurcation.

In this paper, we want to present rumor spreading in a Chinese OSN, sina weibo. We give a thorough analysis by deducing the main-field theory to testify our methods. Our theoretical analyses and simulations show that there is a threshold of rumor spreading in OSN.

The rest of this paper is organized as follows. Section 2 describes the mean-field theory. Section 3 reviews the characteristics of OSN. Section 4 describes the definition of the SIRS model in OSN by using the mean-field theory. Our methods for solution in threshold value of the SIRS model, and simulations and results are discussed in section 5. Finally, Section 6 provides conclusion remarks and the research direction in the future.

2. OSN Analysis with the BA Model

OSN topology follows the Barabasi-Albert (BA) scale-free networks model [15]. Our model is a better alternative to the BA model [1] with extending growth simulation for a Chinese famous OSN, Sina Weibo, which has more than 400 million users. As a conceptual model, the BA model can be applied in many different areas such as biological studies or communication studies. It is a mathematical model where the nodes and links can be defined later so that the model can be applied in diverse areas. The foundations of the model will be introduced in this section. Figure 1 exhibits the scale-free property. A few gray nodes have high degree while all other nodes have low degree

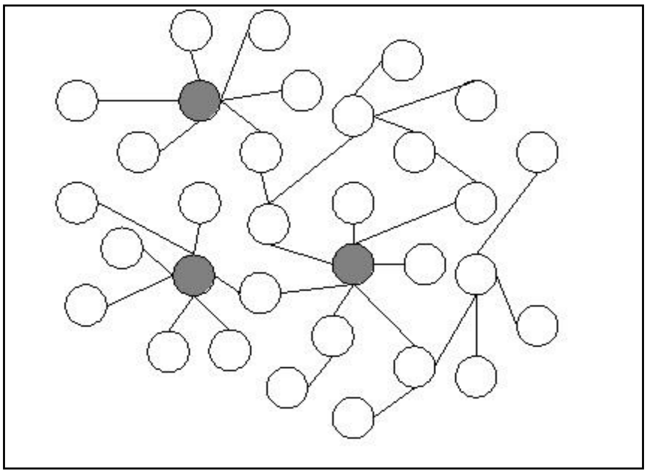

Figure 1. The Graph of the Scale-Free Property

In the BA model, the concept of preferential attachment is followed and the probability $\prod_i(t)$ that a new node gets connected to an existing node *i*, is linearly proportional to the degree $x_i(t)$ of node *i* at iteration *t*, where $\prod_i(t) \sim x_i(t)$.

The conceptual model is motivated by this concept of preferential attachment in the BA model, with further augmentations to the BA model so that it is more flexible and can be applied in more diverse ways such as modeling urban growth behavior.

The probability *P(k)* that a node has *k* incoming links follows a power law: $P(k) \sim k^{-\gamma}$. Using data crawled from information in Sina weibo during 1 July 2013 to 1 October 2013, we can draw the relation between P_r and k in Figure 2.

The parameters q_1, q_2 and q_3 fulfills the condition where $q_1 + q_2 + q_3 = 1$ and are assumed to be constant throughout the evolution of the network. Using continuum formulation [1], the time evolution of the expected node degree for a node *i* is given by:

$$x_i(t+1) = x_i(t) + 2_{q1}\frac{1}{N(t)} + 2_{q2}\frac{x_i(t)}{\sum_j x_j(t)} \quad (1)$$

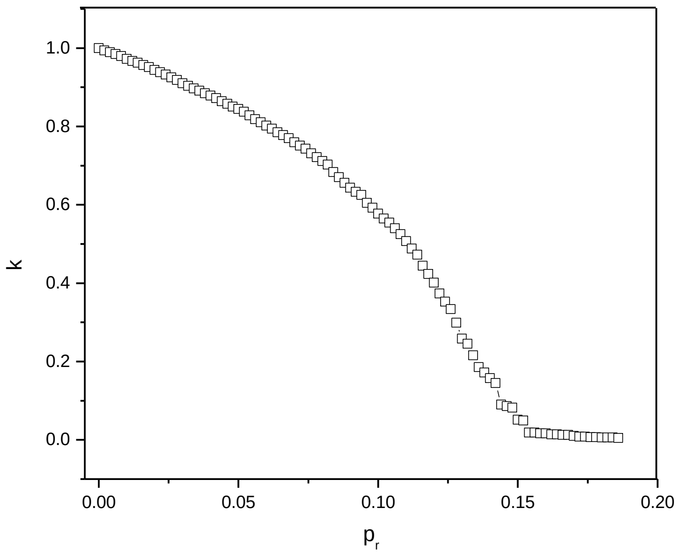

Figure 2. Log-Log Plot of Number of Ratings in Sina Weibo

Where $N(t) = 2\,q_3$ is the expected number of nodes developed after *t* iterations.
After a sufficiently long time, the degree distribution approaches the form:

$$P[x_i = x] \sim (x + B)^{-\gamma} \quad (2)$$

Where $B = \frac{q_1}{q_2 q_3}$,and $\gamma = 1 + \frac{1}{q_2}$, The conceptual model can also reproduce the power law distribution observed in scale free networks when the degrees of node are considered.

After t time steps the network has $N=m_0+t$ nodes and $K=m_t$ edges (ignoring eventual edges of the initial condition). The numerical simulation and the analytical solution of the model in the mean field approximation predicts a degree distribution that asymptotically converges for $t \to \infty$ to

$$P(k) = 2m^2k^{-r} \quad (3)$$

with an exponent $\gamma = 3$, independent of m and of the size N of the network.

3. Rumor Spreading Analysis with Mean-Field Theory

In the BA networks, rumor spreading from the infected hosts to the susceptible ones. Taking a simple model of the spread of an idea, imagine that each person who has heard the idea communicates it with probability μ to each of his or her friends. If the person's degree is k, then the expected number who hear it is μ(k− 1). We make some assumptions in order to concisely and accurately as fowling.

(1)The total number of population N(t) is a variable changing with time t. The online rate b and offline rate μ are constants.

(2)The total population is partitioned into four compartments: the susceptible compartment -spreader (S), the infected compartment - ignorant (I), and the recovered compartment - stifler (R). Each node can only be in a state, and the infected node could not be infected again.

(3)When a node is removed from the infected class, it obtains temporary immunity with probability p and died with probability (1-p).

From the assumptions above, the standard incidence of the total variable population size can be expressed as 1=S(t)+I(t)+R(t).

Table 1 lists parameters and notations needed in this paper.

Table 1. Parameters and Notations in this Paper

Parameter	Notation
S(t)	Number of susceptible nodes (spreader) at time *t*
I(t)	Number of infected nodes (ignorant) at time *t*
R(t)	Number of recovered nodes (stifler) at time *t*
<k>	Average number of links per node
α	Information transmission intensity
β	Infected node cured with probability
γ	Rate of querying wrong knowledge
δ	Probability of node information lossing

The relation of parameters in the paper is showed in Figure 3.

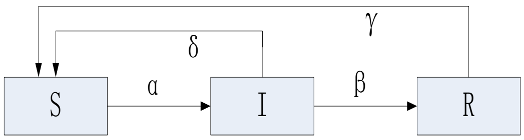

Figure 3. Information Dissemination of SIRS Model in the Paper

So, we get the SIRS model based on mean-field theory as following:

$$\begin{cases} \frac{dS_k(t)}{dt} = -\alpha\langle k\rangle S_k(t)I_k(t) + \delta I_k(t) + \gamma R_k(t) \\ \frac{dI_k(t)}{dt} = \alpha\langle k\rangle S_k(t)I_k(t) - \delta I_k(t) - \beta I_k(t) \\ \frac{dR_k(t)}{dt} = \beta I_k(t) - \gamma R_k(t) \\ S_k(t) + I_k(t) + R_k(t) = 1 \end{cases} \tag{4}$$

The probability of nodes of S condition be infected by nodes of I condition can thus be expressed as $S(t)(1-(1-\lambda)^{<k>I(t)}$.

As for $R_k(\infty) = 1 - S_k(\infty)$ and $R_\infty = 1 - S_\infty$, we get:

$R_\infty = \sum_k P(k)R_k(\infty) = \sum_k P(k)(1 - S_k(\infty)) = 1 - \sum_k P(k)e^{-\lambda k\Phi_\infty} = 1 - S_\infty$

Because $\rho(t)$ is not enough to describe the probability of node, whose degree is k, we use new variable $\Theta(t)$ to express the probability of its. And the expression is:

$$\Theta(t) = \frac{\sum_k kP(k)I_k(t)}{\sum_s sP(s)} = \frac{\sum_k kP(k)I_k(t)}{\langle k\rangle}$$

So, we get density function of uninfected nodes as:

$S_k(t) = e^{-\lambda k R_k(t)/\mu} = e^{-\lambda k\Phi(t)}$

Where, auxiliary function is $\Phi(t) = \int_0^t \Theta(t')dt' = \frac{\sum_k kP(k)R_k(t)}{\langle k\rangle} = \frac{\sum_k kP(k)\int_0^t \rho_k(t')dt'}{\langle k\rangle}$

Then, we get the mean value changing with time as:

$$\frac{d\Phi(t)}{dt} = \frac{\sum_k kP(k)I_k(t)}{\langle k\rangle} = \frac{\sum_k kP(k)(1-R_k(t)-S_k(t))}{\langle k\rangle} = 1-\Phi(t) - \frac{\sum_k kP(k)S_k(t)}{\langle k\rangle} = 1-\Phi(t) - \frac{\sum_k kP(k)e^{-\lambda k\Phi(t)}}{\langle k\rangle}$$

When $t \to \infty$, $\frac{d\Phi_\infty}{dt} = 0$, deduced by above function, we get the equation as:

$$0 = 1 - \Phi_\infty - \frac{\sum_k kP(k)e^{-\lambda k\Phi_\infty}}{\langle k\rangle}$$

Let $f(\Phi_\infty) = 1 - \frac{\sum_k kP(k)e^{-\lambda(k)\Phi_\infty}}{\langle k\rangle} - \Phi_\infty$, because $f(1) < 0$, for $\frac{df(\Phi_\infty)}{d\Phi_\infty}|_{\Phi_\infty=0} > 0$, namely $\frac{\sum_k kP(k)\lambda(k)}{\langle k\rangle} = \lambda\frac{\langle k^2\rangle}{\langle k\rangle} > 1$, that is $\lambda > \frac{\langle k\rangle}{\langle k^2\rangle}$, Φ_∞ must have nonzero solution, as a result, there is a more accurate spread threshold value as $\lambda_c \geq \frac{\langle k\rangle}{\langle k^2\rangle}$.

As for Barabasi-Albert (BA) scale-free model, because $P(k) = \frac{2m^2}{k^3}$, we have:

$$\lambda_c = \frac{\langle k\rangle}{\langle k^2\rangle} = \frac{\sum_k kP(k)}{\sum_k k^2P(k)} = \frac{\sum_k k\frac{2m^2}{k^3}}{\sum_k k^2\frac{2m^2}{k^3}} \approx \frac{\int_m^\infty \frac{2m^2}{k^2}dk}{\int_m^\infty \frac{2m^2}{k}dk} = \frac{2m}{2m^2(\ln\infty - \ln m)}$$
$$\approx \frac{1}{m(\ln N - \ln m)} \sim \frac{1}{\ln N}$$

Corresponding, we get:

$$R_\infty = \sum_k P(k)\left(1-S_k(\infty)\right) = 1 - \sum_k P(k)\,e^{-\lambda k\Phi_\infty} \approx 1 - \int_m^\infty \frac{2m^2}{k^3}e^{-\lambda k\Phi_\infty}dk$$
$$\approx 1 - \int_m^\infty \frac{2m^2}{k^3}\left(\frac{(\lambda k\Phi_\infty)^0}{0!} - \frac{(\lambda k\Phi_\infty)^1}{1!} + \frac{(\lambda k\Phi_\infty)^2}{2!} - \cdots\cdots\right)dk$$
$$\approx 1 - \int_m^\infty \frac{2m^2}{k^3}dk + \int_m^\infty \frac{2m^2}{k^3}\lambda k\Phi_\infty dk$$
$$= 2m\lambda\Phi_\infty$$

For $\Phi(t)$, satisfy $|\lambda k\Phi(t)| < 1$, we can Taylor expand it, and use Γ function ($\Gamma(z) = \int_0^\infty \frac{t^{z-1}}{e^t}dt$) to express it.

$$\frac{1}{\lambda m}\frac{d\Phi(t)}{dt} \approx \Phi(t)(1 - \gamma_E - \frac{1}{\lambda m} - \ln(\lambda m\Phi(t)))$$

We integrate the above function, and get:

$$\Phi(t) \approx \frac{1}{\lambda m}e^{1-\gamma_E-\frac{1}{\lambda m}+Ae^{-\lambda m\Phi(t)}}$$

When $t \to \infty$, $\Phi_\infty \approx \frac{e^{1-\gamma_E}}{\lambda m}e^{-\frac{1}{\lambda m}}$

So, we get $R_\infty \approx 2m\lambda\Phi_\infty \approx 2e^{1-\gamma_E}e^{-\frac{1}{\lambda m}} \sim e^{-\frac{1}{\lambda m}}$.

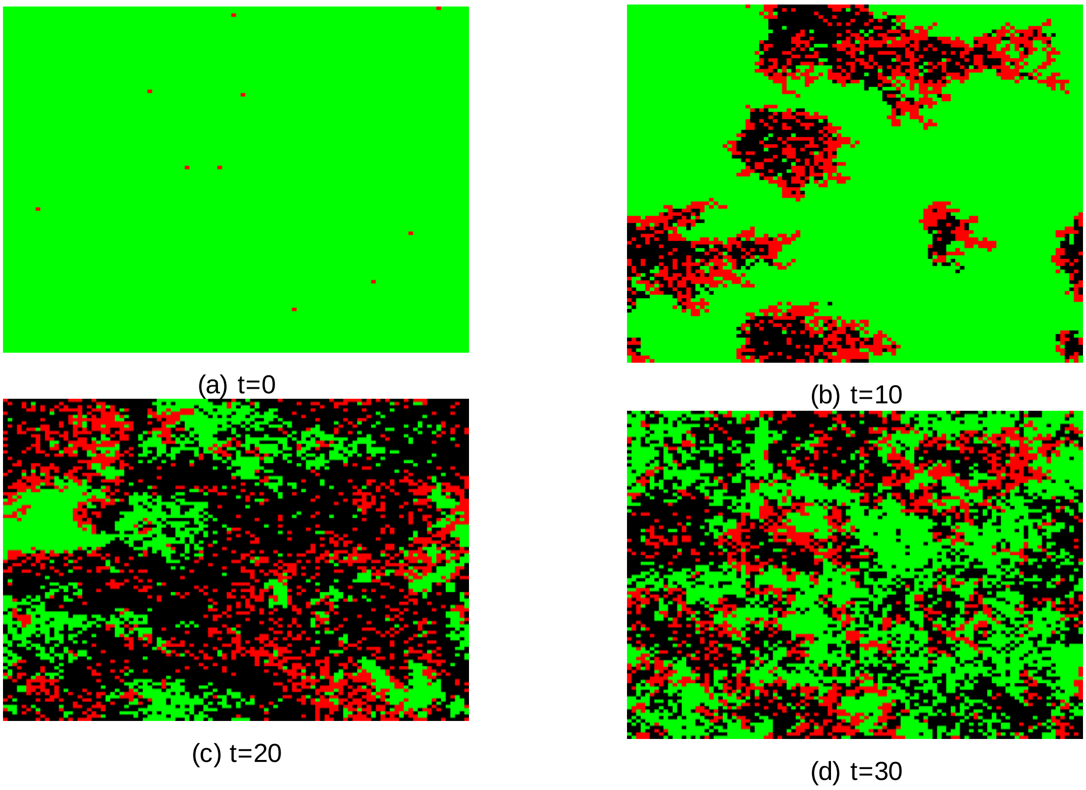

Figure 4. Infected Spots Distribution from (a) to (d) at Time t=0, 10, 20, 30

In Figure 4, we use different color to represent three states in rumor spreading, red is spreader, black is ignorant, and green is stifler. The four panels display evolution of the three states at different time from 0 to 30.

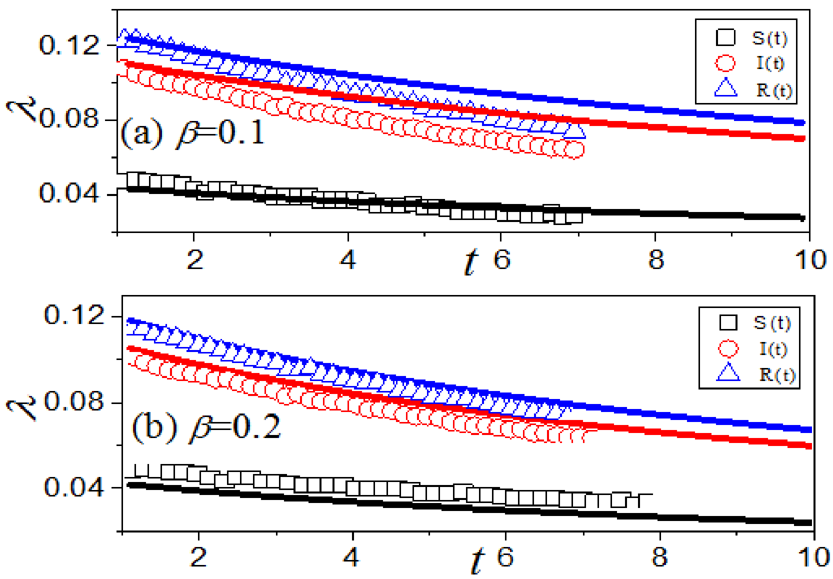

Figure 5. The Three States (S(t), I(t), R(t)) Change Over Time from 0 to 10, Under Different Values (a) β=0.1 and (b) β=0.2

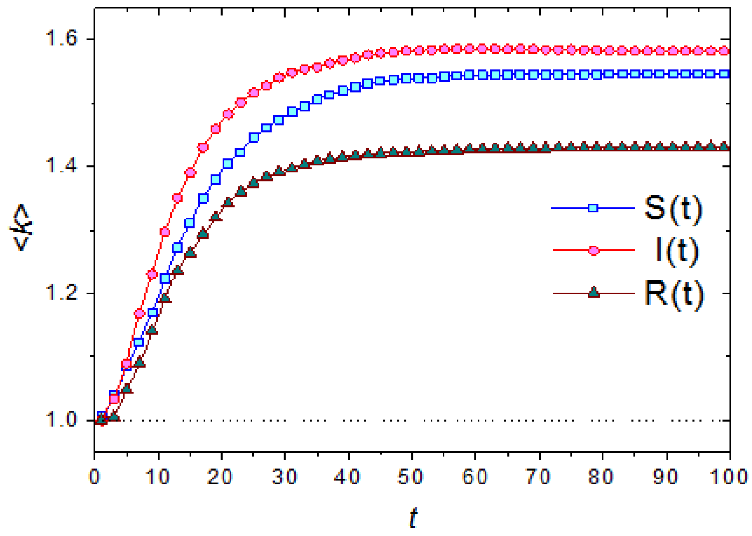

Figure 6. The Average Degree of Networks <k> Changes from Time 0 to 100, Under Three States (S(t), I(t), R(t))

We use Figure 5 to demonstrate that, λ in the range 0~0.14, under different values (a) β=0.1 and (b) β=0.2, the law of three states (S(t), I(t), R(t)) variation over time from 0 to 10.

Figure 6 shows the variation law of three states (S(t), I(t), R(t)) change from time 0 to 100, the relation of time and the average degree of networks <k>. The curve of spreader (S(t), red) is at top of the Figure, which means its average degree <k> is bigger than ignorant (I(t), blue) and stifler (R(t), black).

5. Conclusion

In this paper, we use the dynamics of BA networks approach to study rumor spreading in Sina Weibo. We build scale-free network according to the actual situation of the rumor spreading with mean-field theory. Simulation results are similar with the actual spread trend of information, which verify the validity of the model, proving OSN as an important tool to play a major role on study of rumor spreading.

We illustrate the method to recover scale-free network characteristics using mean-field theory and offer experiments to confirm them. The finding by itself is interesting since it suggests that the informative prior constructed solely upon a theoretical micro-level adoption model increases predictive accuracy of a rumor forecasting model.

Although the scale-free information can be fitted with the SIRS model with a good precision, there is a difference in the parameters of the two information curves. As for the effect of the inhomogeneity of the scale-free networks on the information, we can see it clearly that there is a dependency on the degree of the node that initiates the information, and this dependency goes beyond the evident, namely that a larger number of nodes initiate a larger information. In fact, after a certain node size the information does not grow any further but reaches a finite limit.

Future research directions are both for theoretical as well as experimental works. Phase transition phenomena may be a good examination in this domain, and the spreading thresholds of rumor are exactly to be investigated. Another promising direction is to enhance the model with infected ratio parameter to model environmental factors that affect viral spreading.

Acknowledgment

The work was sponsored by the National Natural Science Foundation of China (Grant No. 71561013), the MOE (Ministry of Education in China) Youth Fund of Humanities and Social Sciences (Grant No. 15YJCZH197), the Jiangxi Provincial Academy of social sciences project (Grant No. 14TQ07), and the Research Project of Science and Technology of Jiangxi Province (Grant No. GJJ150791).

References

[1] L. Barabási and A. Réka, "Emergence of Scaling in Random Networks", Science, vol. 286, no. 5439, **(1999)**, pp. 509-512.

[2] S. Boeealetti, V. Latora, Y. Moreno, M. Chavez and D. U. Hwang, "Complex networks: Structure and dynamics", Physics Reports, vol. 424, **(2006)**, pp. 175-308.

[3] A. E. Motter and Y. C. Lai, "Cascade-based attacks on complex networks", Physical Review E, vol. 66, no. 6, **(2002)**, pp. 65-102.

[4] F. Amblard and G. Deffuant, "The role of network topology on extremism propagation with the relative agreement opinion dynamics", Physical A: Statistical Mechanics and its Applications, vol. 343, **(2004)**, pp. 725-738.

[5] R. P. Satorras and A. Vespignani, "Information dynamics and endemic states in complex networks", Physical Review E, vol. 63, **(2001)**, pp. 71-78.

[6] R. P. Satorras and A. Vespignani, "Epidemic spreading in scale-free networks", Physical Review Letters, vol. 86, no. 14, **(2001)**, pp. 3200-3203.

[7] X. Li and X. Wang, "Controlling the Spreading in Small-World Evolving Networks: Stability, Oscillation, and Topology", IEEE Transactions on Automatic Control, vol. 51, no. 3, **(2006)**, pp. 534-540.

[8] Z. Dezo and A. L. Barabasi, "Halting viruses in scale-free networks", Physical Review E, no. 65, **(2002)**, pp. 31-34.

[9] J. Satsuma, R. Willox, A. Ramani, B. Grammaticos and A. S. Carstea, "Extending the SIRS Epidemic Model", Physical A: Statistical and Theoretical Physics, vol. 336, **(2004)**, pp. 369-375.

[10] A. Barrat, M. Barthélemy and A. Vespignani, "Dynamical Processes on Complex Networks", Cambridge University Press, USA, **(2008)**.

[11] S. Mossa, M. Barthelemy, H. E. Stanley and L. A. N. Amaral, "Truncation of Power Law Behavior in "Scale-Free" Network Models due to Information Filtering", Physical Review Letters, vol. 88, no. 13, **(2002)**, pp. 138-701.

[12] B. K. Mishra and A. Prajapati, "Cyber Warfare: Worms' Transmission Model", International Journal of Advanced Science and Technology, vol. 63, **(2014)**, pp. 83-94.

[13] T. Fan, Y. Li and F. Gao, "Study of Virus Propagation Model in Cloud Environment", International Journal of Security and Its Applications, vol. 7, no. 4, **(2013)**, pp. 267-276.

[14] N. Keshri and B. K. Mishra, "Optimal Control Model for Attack of Worms in Wireless Sensor Network", International Journal of Grid Distribution Computing, vol. 7, no. 3, **(2014)**, pp. 251-272.

[15] A. L. Barabásia, H. Jeonga, Z. Nédaa, E. Ravasza, A. Schubertd and T. Vicsekb, "Evolution of the social network of scientific collaborations", Physical A: Statistical Mechanics and its Applications, vol. 311, no. 3-4, **(2002)**, pp. 590–614.

