A Model of Information Propagation in Social Networks

Jomar F. Rabajante
Institute of Mathematical Sciences and Physics
University of the Philippines Los Baños
College, Laguna, Philippines
+63 49 5636610
jfrabajante@up.edu.ph

Yuichi T. Otsuka
Institute of Mathematical Sciences and Physics
University of the Philippines Los Baños
College, Laguna, Philippines
+63 49 5636610
yuichi0520otsuka@yahoo.com

ABSTRACT
In this paper, we developed a theoretical model to mimic the behavior of the spread of information in social networks. The diversity of social groups can be incorporated in the model, making the model applicable not only to homogeneous networks but also to heterogeneous networks. We started by representing the different states of the actors in the information propagation process using compartments, and then we determined the direct and indirect feedbacks. The model is temporal but the spatial aspect can be considered by grouping the actors based on cluster distances.

The compartment model is then converted into a system of ordinary differential equations (ODEs). We gave a prototype example to simulate the model. The system of ODEs is then solved and analyzed numerically using Berkeley Madonna. The main output is a set of time series on the number of actors present for each state. We also checked the sensitivity of the system to initial conditions and to the given parameters. Such parameters would help us determine the possible optimal spots to control the information propagation.

Categories and Subject Descriptors
J.4 [**Computer Applications**]: Social and Behavioral Sciences – *Sociology;* H.m [**Information Systems**]: Miscellaneous

General Terms
Human Factors

Keywords
Information propagation, compartment model, dynamical systems, differential equations, social contagion

1. INTRODUCTION
It is impractical to have a “good” experiment on the spread of information on the internet because of the diversity and dynamic nature of actors involved. It is even hard to track the whereabouts of the disseminated information. However, this difficulty does not hinder modelers in mimicking the behavior of information flows in complex networks. Some used theoretical models with ordinary differential equations [2, 19]. Others used Markov Chains, Monte Carlo simulations and social network analysis [2, 11, 17, 19].

8th National Conference on Information Technology Education (NCITE 2010), October 20–23, 2010, Boracay, Aklan, Philippines.

Information can be rumors, gossips, formal reports, stories, testimonies, announcements or any other statements; but the most prolific researches are in the study of rumor contagion [4, 10]. Our study follows this research path since the behavior of rumor spread is quite intriguing and fascinating, as well as naturally spontaneous and dynamic.

Rumors are unconfirmed statements [16] passed on from a source to a receiver. Specifically, rumors take place when no clear link exists between people and the correct information, causing ambiguity [1]. When people fail to find a plausible answer to their queries, they begin to interpret the situation and make use of the information at hand to come up with stories [1]. In contrast, gossips are small talks with “inner-circleness” about it and are common in socialization [17].

The basic law of rumor, as formulated by Rosnow and Foster [17], states that rumor strength is directly proportional to the significance of the subject towards the individual concerned and to the uncertainty of the evidence at hand. A modified theory views rumor-mongering as a way of handling anxieties and uncertainties by creating and passing on stories, at times fallacious in nature, attempting to provide an explanation for behavior and to address confusion [15, 16, 17].

In rumor propagation, information is best transmitted by interpersonal interaction [19]. Language, gender and social class also play an important role. Language barriers may serve as hindrances to a rumor, as rumors are another form of communication. Information is not usually transferred from opposite sexes, unless those concerned are related by blood; and the normal trend for rumors is to be confined in a specific social class [18]. Belief in a rumor depends on the degree of suggestibility and credulity of the rumormongers involved.

Two basic models on rumor propagation, the Daley-Kendall [4] and Maki-Thompson Models [10], divide the actors in the rumor process into three mutually exclusive groups called the susceptibles (or ignorants), spreaders, and the stiflers. Susceptibles are those who are not presently aware of a rumor. Spreaders are those who came in contact with the rumor and who are continuously spreading the information. Stiflers are once spreaders but have stopped spreading the information. Variants of these two models have also been developed to model more complicated situations [5, 12, 13, 14]. The Daley-Kendall model is based on the SIR Epidemics model of Kermack and McKendrick [7, 19].

Generally, social network sites possess the ability to spread information, sometimes resulting in the creation of conversations involving various populations. Factors affecting the spread of information in the internet were studied by Magnani, Montesi and Rossi [9]. Another study focused on the activities of blogging, and how information flows through the complex system of blogs [8].

One research analyzed the information propagation on Flickr, an image-centered social network. This study measured the relative speed and the extent of information propagation, and the interaction of the social networks and real-life networks [3]. Moreover, another study focused on information dissemination on mobile networks and produced an efficient algorithm on identifying events leading to information propagation [6].

In this paper, we developed a general model using system of ordinary differential equations (ODEs) for approximating the behavior of the spread of information in social networks. The diversity of social groups can be incorporated in the model. The model is temporal but the spatial aspect can be integrated by extending the grouping of the actors to include cluster distances.

2. GENERAL MODEL

2.1 Compartment Model

The state variables used in our system model are as follows:

For $i = 1,2,\dots,n$ (where n is the number of mutually disjoint groups in the community),

- P_i ~ the set of ignorants in group i;
- P_icS ~ the set of ignorants in group i having contact with spreaders;
- P_iB ~ the set of individuals in group i believing the information; and
- P_iNB ~ the set of individuals in group i not believing the information.

Let $[\phi](t)$ denote the number of individuals present in set ϕ at time t. The compartmental model shown in Figure 1 represents our model.

The groups can be any mutually disjoint factions, clusters or cliques representing all the members in the community. The actors present in the system control the flow of information and thus, are vital in the formulation of our model. As our assumption, we only consider one kind of information unaltered during the process. Other assumptions include: (1) the spread of information is via interpersonal communication (with negligible influence of media broadcasting); (2) the process is deterministic and continuous (i.e. neither stochastic nor discrete); (3) the belief in rumors is affected by the nature of the groups (e.g. socio-economic status, culture, trust system, etc.) [1, 18]; and (4) rates of flow from one state to another are time-homogeneous and follow the law of mass action.

The parameters of the rate of flow from one state to another are represented by the following notations:

For $i = 1,2,\dots,n$,

- a_i ~ the constant average number of new members of group i per unit of time (t);
- b_i ~ the expected proportion of P_i exiting the whole system per unit of time;
- c_i ~ the expected proportion of ignorants in group i having contact with spreaders from the different groups per unit of time;
- d_i ~ the probability that a person from P_icS will believe the conveyed information such that $d_i + e_i = 1$;
- e_i ~ the probability that a person from P_icS will not believe the conveyed information such that $d_i + e_i = 1$;
- f_i ~ the expected proportion of P_iB exiting the whole system per unit of time;
- g_i ~ the expected proportion of P_iNB exiting the whole system per unit of time; and
- h_i ~ the expected proportion of P_iNB returning to P_i to become susceptibles again per unit of time;

When it is infeasible to determine the values of b_i, f_i and g_i, then we can set them to be approximately equal to the average per capita unsubscription rate (u) in the social network, i.e., $1/u$ is the average duration of membership. Note that if $\Delta t \ll 1$, we assumed that an ignorant listens to at most one spreader. The contact rate c_i is affected by the value of $[P_kB], k = 1,2,\dots,n$, as shown by the dashed lines in Figure 1 – i.e., c_i is proportional to the number of believers present. Suppose that $\Delta t \to 0$, the rate c_i can be computed as follows:

$$c_i = \sum_{k=1}^{n} (prob_{ik}) S_k V_{ik} \quad , i = 1,2,\dots,n$$

where

- $prob_{ik}$ ~ probability that an ignorant from group i will have a contact with one spreader from group k per unit of time;
- S_k ~ average proportion of P_kB that are spreaders at time t;
- V_{ik} ~ average number of distinct members of group i listening to a spreader from group k per unit of time; and
- $\sum_{k=1}^{n} prob_{ik} = 1$.

A person can have contact to a spreader via chatting, reading blogs and wall posts, or any other medium of communication in the social network. The rate V_{ik} can also be interpreted as the average number of visits done by distinct members of group i to the social network webpages of a spreader from group k per unit of time. If there are individuals under P_iNB who do not believe the rumor permanently, then we can create a dummy compartment connected to P_iNB that would contain those individuals. A person in P_iB permanently believes the information, and spreaders only come from the set P_iB.

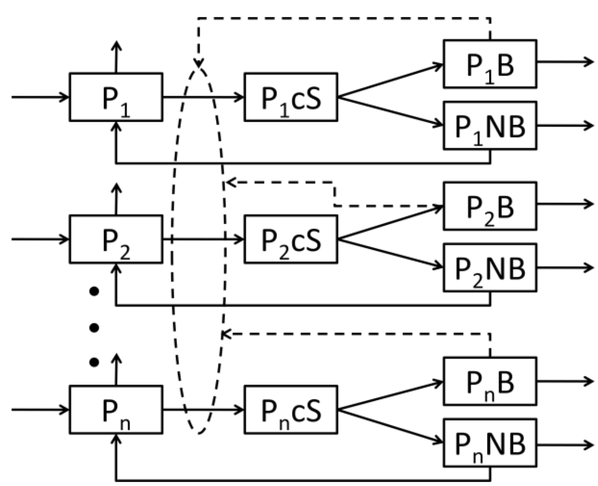

Figure 1. Information Propagation Compartment Model

2.2 Corresponding System of ODEs

The $P_i cS$ are transient states and ${}^{d[P_i cS]}/_{dt} \approx 0,\ \forall\, i = 1,2,\dots,n,$ since it can be presumed that a person usually decides to believe or not to believe an information immediately once the information is conveyed. The system of ODEs that would represent the compartmental model is as follows, $i = 1,2,\dots,n$:

$$\begin{cases} {}^{d[P_i]}/_{dt} = a_i - (b_i + c_i)[P_i] + h_i[P_i NB] \\ {}^{d[P_i B]}/_{dt} = d_i c_i [P_i] - f_i [P_i B] \\ {}^{d[P_i NB]}/_{dt} = e_i c_i [P_i] - (g_i + h_i)[P_i NB] \end{cases}$$

Notice, because of the nature of c_i, that this autonomous and non-homogeneous system of ODEs is nonlinear and coupled.

3. PROTOTYPE EXAMPLES

3.1 Example 1: Closed System

For our example, let us consider two groups. The given parameters and initial values are shown in Figure 2 and Table 1.

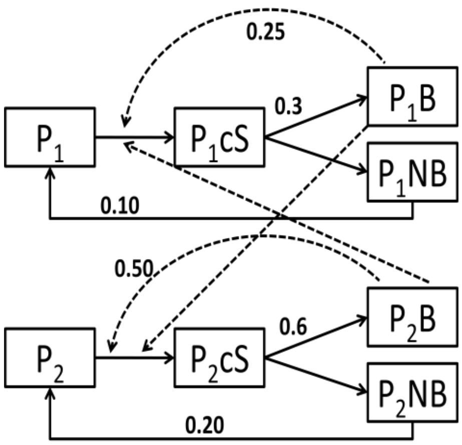

Figure 2. Compartment Model for Example 1

Table 1. Parameters and initial values for Example 1

Para-meter	$prob_{11}$	$prob_{12}$	$prob_{21}$	$prob_{22}$	V_{ik} $(i,k = 1,2)$
	5/1000	3/1000	1/500	6/500	1
Init. Value	$P_1(0)$	$P_2(0)$	$P_1B(0)$	$P_1NB(0), P_2B(0), P_2NB(0)$	
	1000	500	1	0	

The resulting values of c_1 and c_2 are

$$c_1 = \frac{5}{1000}0.25[P_1B] + \frac{3}{1000}0.5[P_2B]$$
$$c_2 = \frac{1}{500}0.25[P_1B] + \frac{6}{500}0.5[P_2B]$$

The corresponding system of nonlinear ODEs is as follows:

$${}^{d[P_1]}/_{dt} = -\left(\frac{1}{800}[P_1B] + \frac{3}{2000}[P_2B]\right)[P_1] + 0.1[P_1NB]$$
$${}^{d[P_1B]}/_{dt} = 0.3\left(\frac{1}{800}[P_1B] + \frac{3}{2000}[P_2B]\right)[P_1]$$
$${}^{d[P_1NB]}/_{dt} = 0.7\left(\frac{1}{800}[P_1B] + \frac{3}{2000}[P_2B]\right)[P_1] - 0.1[P_1NB]$$
$${}^{d[P_2]}/_{dt} = -\left(\frac{1}{2000}[P_1B] + \frac{3}{500}[P_2B]\right)[P_2] + 0.2[P_2NB]$$
$${}^{d[P_2B]}/_{dt} = 0.6\left(\frac{1}{2000}[P_1B] + \frac{3}{500}[P_2B]\right)[P_2]$$
$${}^{d[P_2NB]}/_{dt} = 0.4\left(\frac{1}{2000}[P_1B] + \frac{3}{500}[P_2B]\right)[P_2] - 0.2[P_2NB]$$

Berkeley Madonna [20] was used to solve the system of ODEs. The output time series is shown in the following figures.

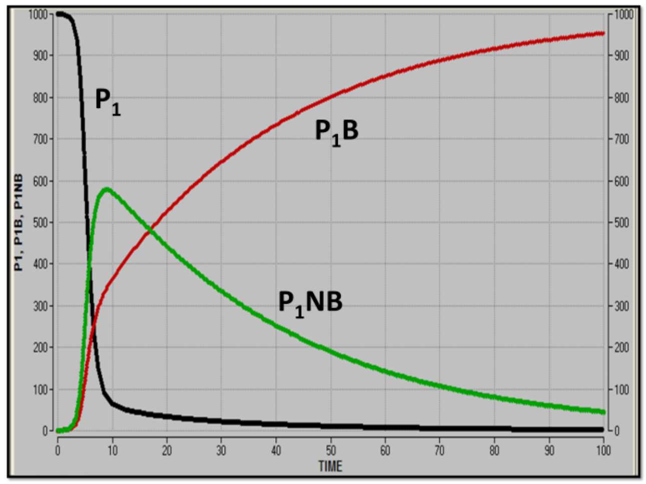

Figure 3. Resulting Time Series for Example 1 (Group 1)

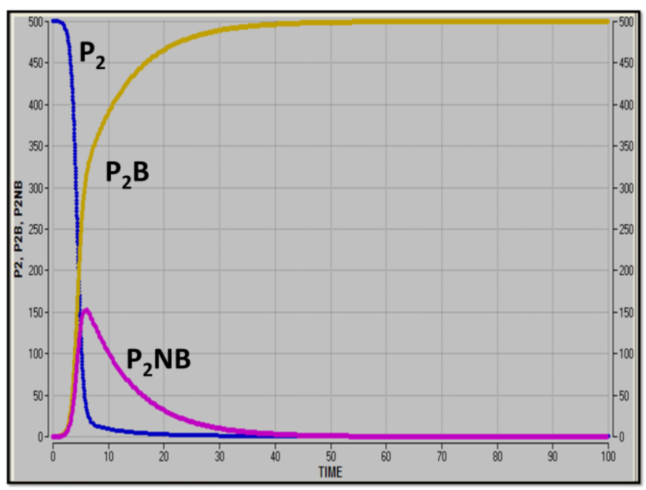

Figure 4. Resulting Time Series for Example 1 (Group 2)

The two figures show two separate graphs of $[P_1]$, $[P_1B]$, $[P_1NB]$ and $[P_2]$, $[P_2B]$, $[P_2NB]$, with respect to time. It can be noted that the number of ignorants ($[P_1]$ and $[P_2]$) continually decreases as time passes. In addition, the number of believers ($[P_1B]$ and $[P_2B]$) continues to increase; while the number of nonbelievers increased to a certain point, but will continuously decrease with respect to time. The system is assumed to be closed, and thus with repeated interactions, the stable equilibrium conditions are $[P_1B](\infty) =$ *total population of group 1* and $[P_2B](\infty) =$ *total population of group 2*, respectively. Thus, P_1B and P_2B are sink compartments.

3.2 Example 2: Open System

For this example, let us consider the situation in Example 1 but with inflows and outflows to and from the system (see Table 2).

Table 2. System Inflow and Outflow Parameters

a_1	a_2	b_1	b_2	f_1	f_2	g_1	g_2
1	2	0.1	0.2	0.1	0.4	0.3	0.2

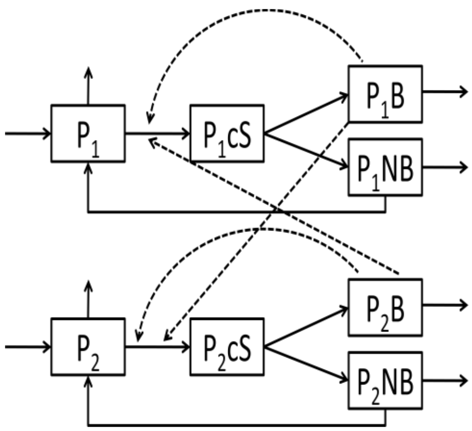

Figure 5. Compartment Model for Example 2

The following set of time series shows the behavior of the increase and decrease in the number of believers and non-believers.

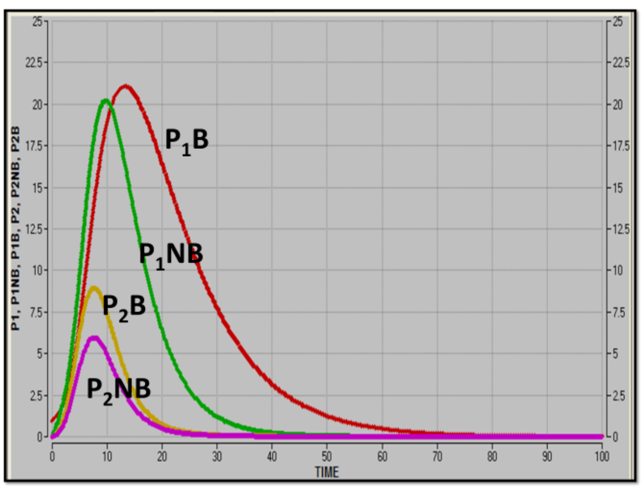

Figure 6. Resulting Time Series for Example 2 (Number of Believers and Non-Believers)

Compared to a closed system, information, such as rumors, may die out in the system because of the inflow of new members and outflow of old members that may have believed the information. The time series may show the duration when a certain information is "hot" which is represented by the neighborhood of the peaks in the value of $[P_1B]$ and $[P_2B]$.

3.3 Sensitivity Analysis and Control

Using Berkeley Madonna [20] we can determine the parameters and initial values where the system is most sensitive to. Determining such parameters would give us some idea in determining what part of the system we can manipulate to control the spread of information.

In the closed system (Example 1), Figure 7 and Figure 8 show the sensitivity of $[P_1B]$ and $[P_2B]$ to initial values. It can be observed that both $[P_1B]$ and $[P_2B]$ are affected by initial values of $[P_1B]$ and $[P_2B]$. However, $[P_1B]$ and $[P_2B]$ are more sensitive to changes in the initial values of $[P_2B]$. This in turn is attributed to the relatively higher probabilities given to $prob_{21}$ and $prob_{22}$.

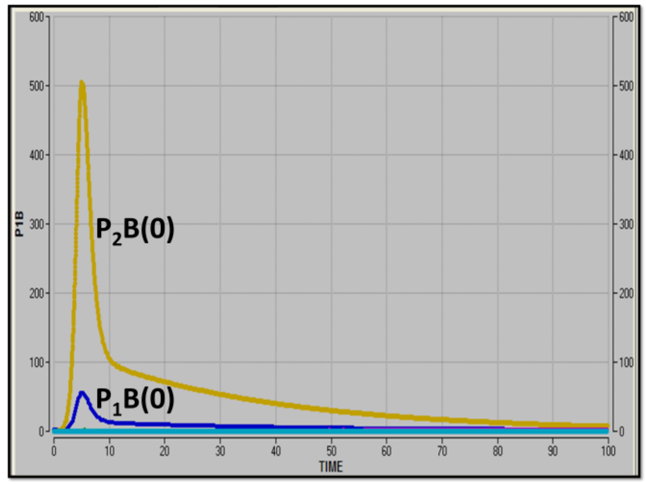

Figure 7. Sensitivity of $[P_1B]$ to initial values (closed system)

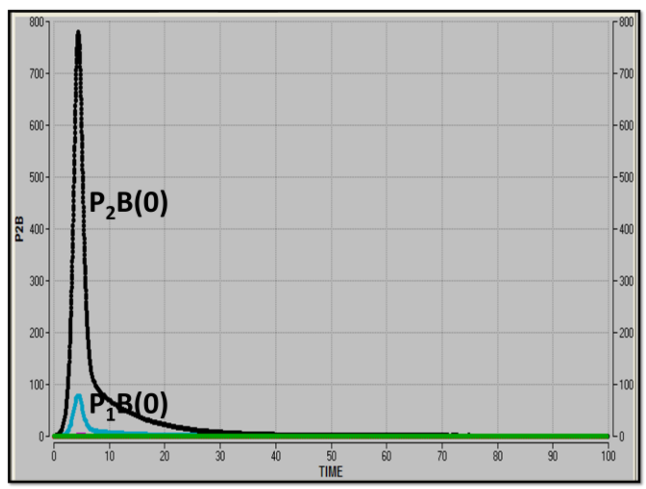

Figure 8. Sensitivity of $[P_2B]$ to initial values (closed system)

Parameter sensitivity analysis can also be done. As an example, we look on what parameters the value of $[P_1B]$ in the open system (Example 2) is most sensitive to (see Figure 9). Notice that $[P_1B]$ is most sensitive to the probability that an ignorant from group i will have a contact with one spreader from group k per unit of time $(i, k = 1,2)$. This means that if we want to control the spread of rumor, we may first look at the feasibility to change these probabilities because these have a great effect in the flow of people to P_1B.

To control the flow of information, such as rumors, we may partition the system into four parts (see Figure 10): (1) inflow of people from the outside of the system, (2) outflow of people from the system, (3) communication point, and (4) belief point. Inflow of people from the outside of the system may represent the new subscriptions of new members of the social network; while outflow may mean unsubscription and delisting of members. A possible effective but usually infeasible way of controlling rumor is delisting the spreaders. We may also control the flow of information through the communication point by regulating the interaction of the members. If not, we may perturb the belief point of members by counteracting the rumor.

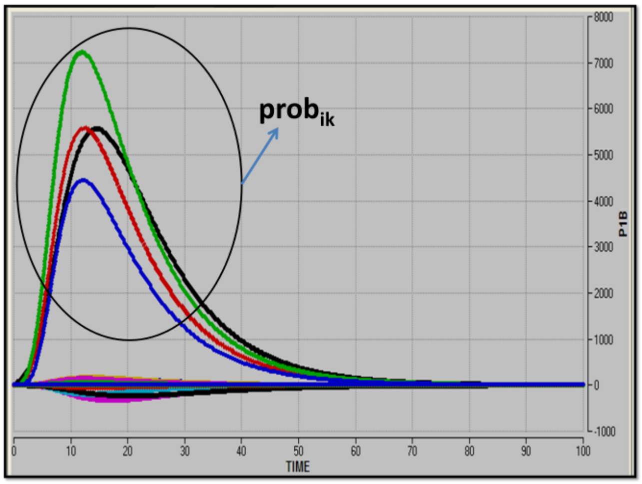

Figure 9. Sensitivity of $[P_1B]$ to parameters (open system)

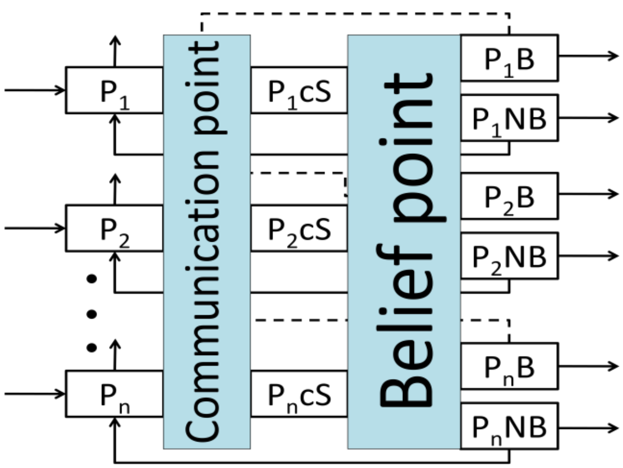

Figure 10. Two of the possible control spots: Communication and Belief points

4. CONCLUDING REMARKS

Even without performing an actual experiment to study information propagation, theoretical modeling helped us understand the behavior of the propagation process. Through this research, we determined the possible control spots. In addition, we were be able to assess the parameters to be manipulated – whether it is more effective to control the inflow or outflow of people to and from the system, or to control the communication schemes or the belief schemes of the actors.

For the social scientists, they can study the details of the control process. In order to regulate the flow of information, the parameter where the system is most sensitive to can be broken down into its elements. Consequently, the social scientists can further analyze the best strategies or tactics for most favorable control. For example, they may look on the factors that affect the belief of the people, and then investigate the degree on how these factors contribute to the flow of information.

Future research may extend this study to include the war of ideas and multiple conveyed information. We may also try to model the discrete and stochastic nature of the system. However, this model would not be useful if the parameters were not measured in accordance to reality.

5. ACKNOWLEDGMENTS

Our thanks to the College of Arts and Sciences and to the Institute of Mathematical Sciences and Physics, UPLB for the financial support. This research would not be possible without the motivation and inspiration coming from the collaboration with the Department of Social Sciences, UPLB, especially to: Prof. Girlie Nora A. Abrigo, Prof. Gloria Luz M. Nelson and Prof. Joyce A. Ocampo.

6. REFERENCES

[1] Bordia, P. and DiFonzo, N. 2004. Problem solving in social interactions on the Internet: Rumor as Social Cognition. *Social Psychology Quarterly*. 67 (2004), 33-49.

[2] Brady, K. 2009. *Rumor Models*. Department of Mathematical Sciences Carnegie Mellon University. DOI=http://www.math.cmu.edu/~howell4/teaching/bradyproj.pdf

[3] Cha, M., Mislove, A., and Gummadi, K. P. 2009. A measurement-driven analysis of information propagation in the flickr social network. In *International World Wide Web Conference* (Madrid, Spain, April 20-24, 2009). WWW 2009. 721-730.

[4] Daley, D. and Kendall, D. 1965. Stochastic Rumours. *Journal of Applied Probability*. (1965), 42-55.

[5] Dodds, P. and Watts, D. 2004. A generalized model of social and biological contagion. *Journal of Theoretical Biology*. (2004) 587-604.

[6] Dyagilev, K., Mannor, S., and Elad, Y-T. 2010. Generative Models for Rapid Information Propagation. In *1st Workshop on Social Media Analytics* (Washington, D.C., July 25, 2010). SOMA 2010. ACM, New York, 1-9.

[7] Kermack, W. and McKendrick, A. 1927. A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London*. (1927), 700-721.

[8] Leskovec, J., McGlohon, M., Faloutsos, C., Glance, N., and Hurst, M. 2007. Information Propagation and Network Evolution on the Web. *Machine Learning Department - Carnegie Mellon University*. DOI= http://www.ml.cmu.edu/current_students/McGlohonKDD.pdf.

[9] Magnani, M., Montesi, D., and Luca, R. 2010. Information propagation analysis in a social network. In *The 2010 International Conference on Advances in Social Networks Analysis and Mining* (Odense, Denmark). 1-5.

[10] Maki, D. and Thompson, M. 1973. *Mathematical models and applications, with emphasis on the social, life, and management sciences*. Prentice Hall, Englewood Cliffs, N.J.

[11] Moreno, Y., Nekovee, M., and Pacheco, A. F. 2004. Dynamics of rumor spreading in complex networks. *The American Physical Society*. (2004).

[12] Nekovee, M., Moreno, Y., Bianconi, G., and Marsili, M. 2006. Theory of rumor spreading in complex social networks. *Physica A*. (2006), 457-470.

[13] Pezzo, M. V. and Beckstead, J. W. 2006. A Multilevel Analysis of Rumor Transmission: Effects of Anxiety and Belief in Two Field Experiments. *Basic and Applied Science Psychology.* (2006), 91-100.

[14] Piquiera, J. R. 2010. Rumor propagation model: an equilibrium study. *Mathematical Problems in Engineering*. (2010), 1-7.

[15] Rosnow, R. 1991. Inside rumor: A personal journey. *American Pyschologist.* (1991), 484-496.

[16] Rosnow, R. L. 2001. Rumor and gossip in interpersonal interaction and beyond: A social exchange perspective. *American Psychological Association.* (2001), 203-232.

[17] Rosnow, R. L. and Foster, E. K. 2005. Rumor and Gossip Research. *Psychological Science Agenda*. 19, (April. 2005), 1-2.

[18] Scanlon, J. T. 1977. Post-disaster Rumor Chains: A Case Study. *Mass Emergencies*. 121-126.

[19] Schramm, H. C. 2006. *An Analytic Framework for the War of Ideas.* Master's Thesis. Naval Postgraduate School.

[20] Berkeley Madonna (ver. 8.3.18). Macey R. I. and Oster G. F. (© 1996-2010). University of California.