2012 32nd International Conference on Distributed Computing Systems Workshops

Diffusive Logistic Model Towards Predicting Information Diffusion in Online Social Networks

Feng Wang, Haiyan Wang, Kuai Xu
Arizona State University
Email: {fwang25, haiyan.wang, kuai.xu}@asu.edu

***Abstract*—Online social networks have recently become an effective and innovative channel for spreading information and influence among hundreds of millions of end users. Most of prior work either carried out empirical studies or focus on the information diffusion modeling in temporal dimension, little attempt has been given on understanding information diffusion over both temporal and spatial dimensions. In this paper, we propose a Partial Differential Equation (PDE), specifically, a Diffusive Logistic (DL) equation to model the temporal and spatial characteristics of information diffusion. We present the temporal and spatial patterns in a real dataset collected from a social news aggregation site, Digg, and validate the proposed DL equation in terms of predicting the information diffusion process. Our experiment results show that the DL model is able to characterize and predict the process of information propagation in online social networks. For example, for the most popular news with 24,099 votes in Digg, the average prediction accuracy of DL model over all distances during the first 6 hours is 92.08%. To the best of our knowledge, this paper is the first attempt to use PDE-based model to study the information diffusion process in both temporal and spatial dimensions in online social networks.**

I. Introduction

Recent years have witnessed the explosive growth of online social networks that connect people in the digital world. One of the most important functionalities of online social networks is to effectively spread information such as latest news headlines, movie recommendations, etc, across the networks. This process is called information diffusion. Given the significant role online social networks have played in elections and crisis [1], it has become increasingly urgent to gain a deep understanding of information diffusion process through online social networks. However, understanding the process of information diffusion over online social networks is a daunting task due to the intricacy of human dynamics and social interactions, the vast scale of users and information, and complexity of these networks.

Most prior work in information diffusion studied the characteristics of information diffusion over various online social networks using empirical approaches [2], [3], [4], [5]. A few recent effort use mathematical models to predict information diffusion over a time period in online social networks [6], [7]. However, little attempt has been given on understanding and modeling information diffusion in both temporal and spatial dimensions.

This paper explores one key question: *How does a piece of information travel over time and space in an online social network?* Specifically we are interested in answering the *spatio-temporal diffusion problem*: for a given information m initiated from a particular user called *source* s, after a time period t, what is the density of influenced users at distance x from the source? An influenced user is an user who has actively voted or liked the information. We use friendship hops as distance and abstractly translate the information diffusion process in online social networks into two separate processes: growth process and social process. Growth process represents information spreading among users with the same distance from the source and social process is the process through which information randomly spreads among users at different distances from the source.

We carry out empirical studies in a real dataset collected from Digg, a major social news aggregation site. We find out that friendship hops is a good indicator of distance in Digg networks. The densities of influenced users at different distances show consistent evolving pattern. This affirms our choice of a PDE-based DL equation to model information diffusion in online social networks. Furthermore, we validate the proposed DL model on the Digg dataset. The experiment results show that by constructing the proper initial condition and parameters, the DL model can effectively predict the density of influenced users for a given distance and a given time for both distance metrics. For example, for the most popular news with 24,099 votes in Digg, the average prediction accuracy of DL model over all distances during the first six hours is 92.08%.

To the best of our knowledge, this paper is the first attempt to study the spatio-temporal diffusion problem in online social networks and propose PDE-based models for characterizing and predicting the temporal and spatial patterns of information diffusion over online social networks. The contributions of this paper include:

- We introduce the spatio-temporal diffusion problem to understand information diffusion in online social networks;
- We abstract the diffusion process and introduce DL equation to model diffusion process in online social networks;
- We present the temporal and spatial patterns of infor-

1545-0678/12 $26.00 © 2012 IEEE
DOI 10.1109/ICDCSW.2012.16

mation diffusion in a real dataset collected from a major online social news aggregation site;

- We validate the DL model and evaluate its performance by matching its prediction with the real dataset.

The remainder of this paper is organized as follows. Section II introduces the DL model for modeling information diffusion in online social networks, describes the construction of the initial density function. Section III presents empirical studies of the temporal and spatial patterns in the Digg datasets, and validates the effectiveness of the DL model in terms of prediction accuracy. Section IV gives a brief literature review on related work, and Section V concludes the paper and outlines our future work.

II. Diffusive Logistic Model

In this section, we first introduce friendship hops as a distance metric, then describe a PDE-based Diffusive Logistic model to characterize information diffusion process. Subsequently we give guidelines on parameter selection.

A. Distance Metric

Defining distance is a key task to understand spatial diffusion pattern. A good distance metric should capture the information diffusion channel in a high level. Since one major channel of information diffusion is through friends, we adopt *friendship hops* which is the smallest number of hops from one user to another in a social friendship graph, to quantify the distance. It is worth noting that this metric is simple but powerful. For example, it in some extent captures the structure of the network since a user with large node degree in the network topology can impact the distance of a large set of nodes which are his neighbors.

B. Diffusive Logistic Model

Information and influence spread dynamically in many different ways through the complex online social networks. It is challenging to tell the topology of the underlying diffusion networks [8]. In general, the information can not only spread along social links between users that are direct friends but also spread in a random way among users sharing similar interests but without direct social links. We propose a PDE-based Diffusive Logistic model to characterize the information diffusion process. We propose an innovative approach to abstractly translate the diffusion into two processes which can be respectively modeled with mathematical models widely used in mathematical biology, sociology, and physics.

1) Model Heuristic: Let U denote the user population in an online social network, and s is the source of an information. For social network users, based on their distances to the source, the user population U is breakdown into a set of groups, i.e., $U = \{U_1, U_2, ...U_i, ..., U_m\}$, where m is the maximum distance. The group U_x consists of users that have distance x to the source.

As information propagates through social networks, some users express their interest in the information by commenting, voting, forwarding, digging or other activities. We call such users as *influenced users* of the information. Let $I(x, t)$ denote the density of influenced users at distance x during time t, that is, $I(x, t)$ reflects the ratio of the number of influenced users with a distance of x at time t over the total number of users in U_x. The value of $I(x, t)$ depends on two diffusion processes. First, the users in U_x could also influence each other. In online social networks, social triangles, also called triads formed by high clustering of users, are very common in online social networks. Therefore, it is possible that two users of the same distance from the source are also friends themselves. We call this process *growth process*. Secondly, the users in U_y where $y \neq x$ can influence those in U_x through direct or indirect friendship links that can be either uni-directional or bi-directional. We call it *social process*. We argue that this diffusion is in the manner of *random walk*, that is, users at distant x randomly influence the users at a different distance. This is true for social media sites such as Twitter and Digg. In these sites, information spreading happens when a follower retweets/votes for news submitted by his followee. In addition, a user, who is not a follower of the users who have retweeted/voted a news, can also retweets/votes for the same news after the news is promoted to the front page, or listed by the search functions of the sites. Hence, information propagation also randomly happens between two users who are not direct friend.

Figure 1 illustrates the growth process and social process. The circle $x = i$ represents all the users that have distance i from the source. It contains many users that spread information among themselves. Users at a certain distance can spread information to users at other distances in a random fashion.

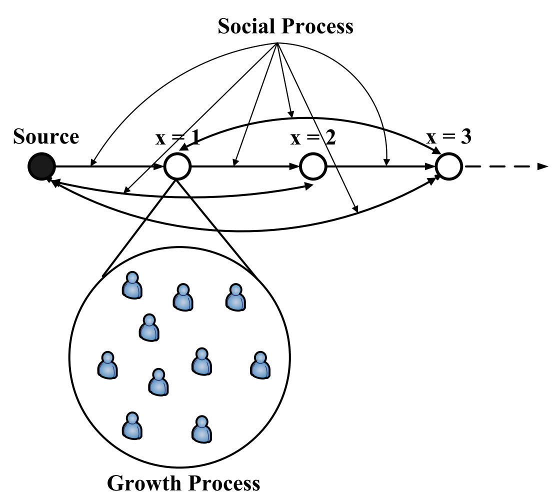


Figure 1. Information spread processes in online social networks

2) Model Description: The growth process can be modeled with logistic model [9]. Logistic model is widely used to model the population dynamics where the rate of reproduction is proportional to both the existing population and the amount of available resources. It has also been used to describe various population dynamics and predict growth of bacteria and tumors over time, etc [9]. In the context of online social networks, we are interested in modeling the impact of the user influence within the same group U_x on the growth of $I(x,t)$, the density of influenced users at the distance x during time t. Hence, the growth process is modeled with logistic equation as:

$$\frac{\partial I(x,t)}{\partial t} = rI(x,t)(1 - \frac{I(x,t)}{K}) \tag{1}$$

The social process is similar to spatial spread of infectious disease in epidemiology [10], thus it can be measured with $d\frac{\partial^2 I}{\partial x^2}$ if the social capability among users at different distances is set as d.

Combining the social process and growth process together, we derive the DL equation as follows:

$$\begin{aligned} &\frac{\partial I}{\partial t} = d\frac{\partial^2 I}{\partial x^2} + rI(1 - \frac{I}{K}) \\ &I(x,1) = \phi(x), \ \ l < x < L \\ &\frac{\partial I}{\partial x}(l,t) = \frac{\partial I}{\partial x}(L,t) = 0, \ \ t > 1 \end{aligned} \tag{2}$$

where

- I (interchangeable with $I(x,t)$) represents the density of influenced users with a distance of x at time t;
- d represents the social capability measuring how fast the information travels across distances in social networks;
- r represents the intrinsic growth rate of influenced users with the same distance, and measures how fast the information spreads within the users with the same distance;
- K represents the carrying capacity, which is the maximum possible density of influenced users at a given distance;
- L and l represent the lower and upper bounds of the distances between the source s and other social network users;
- $\phi(x)$ is the initial density function. It is non-negative and not identical to 0. Each information has unique ϕ which can be constructed from the initial phase of spreading;
- $\frac{\partial I}{\partial t}$ represents the first derivative of I with respect to time t;
- $\frac{\partial^2 I}{\partial x^2}$ represents the second derivative of I with respect to distance x.

$\frac{\partial I}{\partial x}(l,t) = \frac{\partial I}{\partial x}(L,t) = 0$ is the Neumann boundary condition [9], which means no flux of information across the boundaries at $x = l, L$. This is true for online social networks since information spreads within the networks.

C. Parameter Selection

Now we present the method to construct the initial density function ϕ and provide some guidelines for model parameter selection. In general, the initial function is constructed using the data collected from the initial stage of information diffusion. Specifically, the function ϕ which is a function of distance x captures the density of influenced user at distance x at time $t = 1$.

The DL model has three requirements on the initial function ϕ: i) the function has to be twice continuous differentiable; ii) the slopes at the left and right ends are zero, that is, $\phi'(l) = \phi'(L) = 0$; and iii)

$$d\phi'' + r\phi(1 - \frac{\phi}{K}) \geq 0. \tag{3}$$

In online social networks, it is only possible to observe discrete values for the initial density function, because the distance x is discrete. To satisfy the first requirement of the model, we apply a simple and effective mechanism available in Matlab cubic spline package, called *cubic splines interpolation* [11], to interpolate the initial discrete data in constructing $\phi(x)$. Using this process, a series of unique cubic polynomials are fitted between each of the data points, with the stipulation that the obtained curve is continuous and smooth. Hence $\phi(x)$ constructed by the cubic splines interpolation is a piecewise-defined function and twice continuous differentiable. After cubic splines interpolation, we simply set the two ends to be flat to satisfy the second requirement, since in this way the slopes of the density function $\phi(x)$ at the left and right ends are zero. For the last requirement, we have $\phi(x) \leq K$, since K is the carrying capacity. Thus $r\phi(1 - \frac{\phi}{K}) \geq 0$. If ϕ is convex, then $d\phi'' \geq 0$ and Equation 3 holds[1]. If ϕ is concave in some range, Equation 3 will still hold, as long as K is relatively large and the social capability d is sufficiently smaller than growth rate r.

Parameters r, d, K in the DL model can be constants or functions of time t and distance x. In general, growth rate r controls the gap between $I(x,t)$ and $I(x,t{+}1)$ and is usually a function of t. Social capability d controls the slope of I, and carrying capacity K controls the upper bound of I.

III. Experiment Evaluations

In this section, we present our findings of information diffusion characteristics in an anonymized Digg dataset [12] and evaluate the performance of the proposed DL model with the dataset. We'd like to answer two specific questions: first, what is the density of influenced users at distance x from the source at time t; second, how to forecast and predict

[1] In our experiments with Digg data, most region of ϕ is convex using friendship hops as distance.

the density given density data collected at the initial stage of a news. We first describe the anonymized Digg dataset, then present the temporal and spatial characteristics of the dataset, and finally validate the DL model by comparing the predicted density with the actual observations in the dataset.

A. Digg Data Set

The dataset used in this study is collected from Digg, one of the most popular news aggregation sites. Users can submit links of news stories that they find in professional news sites and blogs to Digg, and can vote and comment on the submitted news. Digg users form friendship links through "following" each other. The first voter who brings the news to the Digg site, is called the initiator or source. There are two ways of information propagation in Digg: 1) A user can see the news submitted by the friends he follows and vote the news. After a user votes for a news, all his followers are able to see and vote on the news, and so on. 2) Once the news is promoted to the front page due to high popularity, the users, who do not friend with the initiator directly or indirectly, will also be able to view and vote for the news. A user can also discover news through search engines provided by the site and vote for it. The second approach of information propagation confirms our assumption of random walk of information spreading. Thus the Digg data provides an opportunity for us to study the impact of the friendship relationships and random walk of diffusion on the process of information spreading.

The datasets consist of 3553 news stories that are *voted* (also called *digged*) and promoted to the front page of `www.digg.com` due to vote popularity during June 2009. In total, there are more than 3 millions votes cast on these news stories from over $139,409$ Digg users. In addition, the datasets also include the directed *friendship* links among the Digg users who have voted these news stories. Based on these friendship links, we construct a directed social network graph among these Digg users. For each of the news stories, the datasets include the user id of all the voters during the collection period, and the timestamps when votes are cast. The time granularity is in seconds. The timestamp and the social network graph provide critical data to study the temporal and spatial patterns of information propagation.

B. Characterizing Temporal and Spatial Patterns of Information Diffusion in Digg

To gain a better understanding of information diffusion in online social networks, we characterize the temporal and spatial patterns of information diffusion process in Digg dataset. Specifically we study how the information propagates over *time* in online social networks, and examine the impact of the *distance* measured by *friendship hops*. Since each news is independent, we demonstrate the results of the four representative news. Story $s1$ and $s2$ are the two most popular news in the Digg dataset with 24,099 votes and 8521 votes respectively, story $s3$ is a less popular news with 5988 votes, and story $s4$ is a news with only 1618 votes. These four news are chosen to represent news of different scales, other news of similar scales follow the same propagation characteristics as these four news.

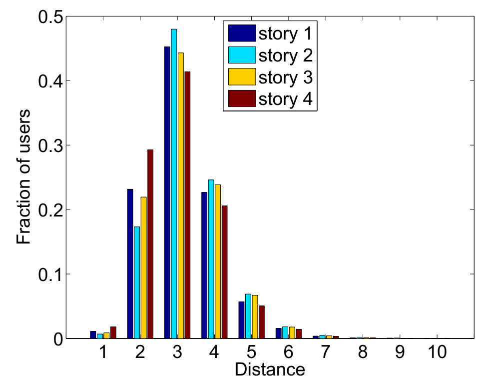


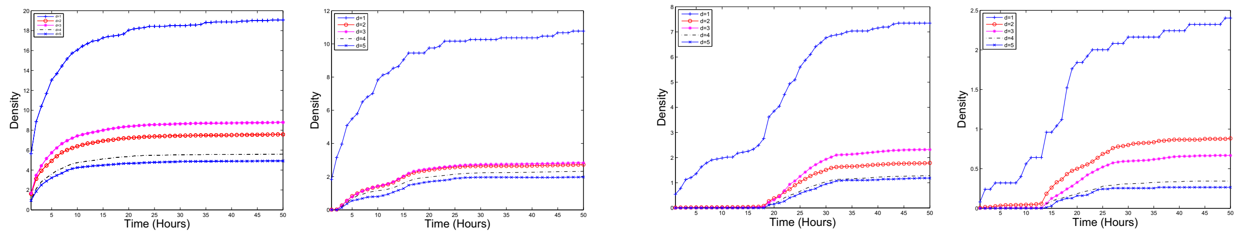
Figure 2. Distribution of neighbors of the initiators of the four representative stories

As discussed in the previous section, we define the distance between the initiator and any other user as the length (the number of friendship hops) of the shorted path from the initiator to this user in the social network graph. Clearly, the direct followers of the initiator have a distance of 1, while their own direct followers have a distance of 2 from the initiator, and so on. Figure 2 shows the distance distributions of the direct and indirect followers of the initiators of the four representative stories. An interesting observation is that the majority of social network users have a distance of 2 to 5 from the initiators. For example, for all four stories, the users of distance 3 account for more than 40% of all the users reachable from the initiators. As the distance increases from 6 to 10, the number of social networks users at these distances drops sharply. Due to the small size of users of distance 6 to 10, we only present the results for users of distance 1 to 5 in the remaining of this section.

Figure 3[a-d] illustrate the density of influenced users at 1 to 5 friendship hops from the initiators over 50 hours for four stories. In the context of Digg social networks, we consider the users who have voted the news story as influenced users. x-axis is the time, y-axis is the density of influenced users. Each line represents the density at a certain distance. The lowest line is the density at distance 1.

There are five interesting observations:

- The densities of influenced users at different distances show consistent evolving pattern instead of increasing randomly. This phenomenon can be described by spatial-temporal PDE, which leads us to the DL model;
- Figure 3[a] shows that for story $s1$, the density of influenced users that are 3 hops away from the initiator is higher than that of users 2 hops away. This verifies



(a) Density of influenced users of s1 (b) Density of influenced users of s2 (c) Density of influenced users of s3 (d) Density of influenced users of s4

Figure 3. Density of influenced users over 50 hours with friendship hops as distance

that social links are not the only channel of information diffusion in Digg networks. Otherwise, if information only spreads along social links, then intuitively, as the distance increases, the density should decrease;

- The density of influenced users at distance 1 (represented by the top line) is significantly higher than that of users with hops greater than 1. For story $s4$, the density of influenced users decreases as hops increases. This indicates that even there are other channels, the social links do play an important role in information diffusion in Digg networks;
- Popular stories spread faster than less popular ones. For example, the density of influenced users of story $s1$ remains stable after about 10 hours while story $s2$ (shown in Figure 3[b]) is stable after about 20 hours;
- A consistent observation over all these sample stories is that after 50 hours, the densities of influenced users for all distance remain stable, which suggests that the news stories are no longer "new" anymore.

C. Predicting Information Diffusion with the DL model

In this subsection, we evaluate the DL model given the initial spreading phase of a story. The construction of initial density function follows the method outlined in previous section. Specifically, we create the initial density function for each news story using the density of influenced users captured at the first hour after the story was initially posted on Digg.

Figure 4 illustrates the predicted results for the most popular story $s1$[2]. In our evaluation, the carrying capacity K is set to 25 and social capability d is set to 0.01. Growth rate is set as a decreasing function of time t and is defined as $r(t) = 1.4e^{-1.5(t-1)} + 0.25$. It is worth noting that even different news stories may need different parameters, the precise prediction of the most popular news in Digg demonstrate the effectiveness of the proposed model. x-axis is the distance measured by *friendship hops* and y-axis represents the density of influenced users. The dashed lines denote the *actual* observations for the density at different times, while the solid lines illustrate the density *predicted* by the DL model. The lowest line representing $t = 1$ is the initial density function. In online social networks, the density is only meaningful when distance is integer. It is clear that the predicted values closely match the actual values over time and distance.

[2] Other four news can be precisely modeled with adjusted parameters. The result is omitted due to space limit

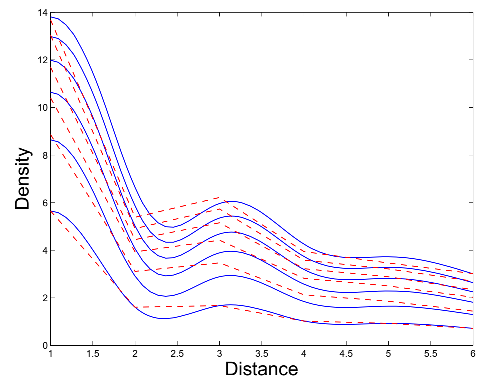


Figure 4. Predicted vs. Actual data of story s1

To quantify the accuracy of the prediction, we calculate the prediction accuracy of the DL model using the following formula: $1-\frac{|\text{predicted value}-\text{actual value}|}{\text{actual value}}$. Table I gives the average accuracy for users of distance 1 to 6 of the most popular story $s1$. It is calculated from the same data illustrated in Figure 4. As we can see, with the setting of initial density function, r, K and d, the DL model gives most accurate prediction for users at distance 1 where the average prediction accuracy is 98.27%. The overall average prediction accuracy across all distances is 92.81%, which indicates the effectiveness of the DL model.

In summary, the experiment results based on Digg datasets show that our PDE-based DL model effectively characterizes and predicts the process of information cascading over online social networks.

IV. Related Work

Information diffusion over online social networks has drawn much attention from the research community [13]. Many prior work have focused on empirical studies in

Table I
THE PREDICTION ACCURACY WITH FRIENDSHIP HOP AS DISTANCES FOR STORY $s1$

Distance	Average	t = 2	t = 3	t = 4	t = 5	t = 6
1	98.27%	97.47%	97.74%	97.48%	99.55%	99.09%
2	86.99%	93.59%	96.63%	87.16%	80.80%	76.78%
3	90.28%	83.23 %	87.98%	90.99%	93.35%	95.94%
4	92.98%	86.75%	91.39%	99.00%	95.68%	92.06%
5	93.77%	89.05%	91.61%	97.79%	97.92%	92.49%
6	94.56%	90.03%	89.48%	96.04%	97.57%	99.67%

different online social networks. For example, [2] examined how the interest in news stories spreads among the users in Digg and Twitter social networks, [3] studied the factors that prohibit the epidemic transmission of popular news posted on Digg, [4] measures the message propagation on Twitter, and Tang *et al.* presented a large-scale empirical study on network structure, user characteristics, and content dissemination process of Digg [14]. In addition, [15] studied the dynamics of information propagation in weblogs. [5], [16], [17] studied cascade patterns of disseminating popular photos over Flickr social network.

Several recent work have proposed mathematical models for understanding and predicting information diffusion in online social networks over time. [18] characterized the process of information diffusion over social networks during a given time period using SIS (Susceptible, Infected, and Susceptible) epidemic model. An earlier work [19] used two most basic diffusion models, namely *Linear Threshold* and *Independent Cascade Models*, in searching the most influential users in online social networks. [6] introduced a Linear Influence Model to predict the number of newly infected nodes based on the time when previous set of nodes are infected. [7] used mathematical models to understand information diffusion in temporal dimension. [20] proposed a model to infer underlying paths of information diffusion, while [21] studied how to limit the spread of misinformation.

In addition, several research work have been done to study the impact of the structure of online social networks on the process of information diffusion. In particularly, [22] studied the information structure of Weblogs and microblogs and revealed their systematical difference in contribution and navigation patterns and user interactions, while [23] examined the impact of human activity patterns on the dynamics of information diffusion using a viral email experiment involving over 31 thousand individuals. Further, [24] demonstrated the effects of users' connection patterns in an online social network on the information diffusion process.

In parallel with research in computer science, there are also research in biology, sociology, economics, and physics to model time evolution systems [25], [26] which use differential equation models to translate local assumptions or data on the movement and reproduction of individuals into global conclusions on the population.

Our work is different from these prior studies in online social networks because we focus on mathematical models to predict the information diffusion in both temporal and spatial dimensions in online social networks, which has not been addressed in any of these studies. Furthermore, we are different from dynamic mathematical modeling in other disciplines since social distance, information diffusion, and influence growth are abstract and unique for online social networks.

V. CONCLUSIONS AND FUTURE WORK

This paper introduces a novel PDE-based DL model for solving the *spatio-temporal diffusion problem* in online social networks. Through measuring the density of influenced users at a certain distance from the source of the information during a given time period, we characterize the temporal and spatial patterns of information diffusion process. Further, we use the proposed DL model to predict the density of influence users with varying distances over time based on the early phase of information spreading activities, and evaluate the prediction quality using datasets collected from Digg social networks. Our future work include: 1) developing new models that consider social capacity, growth rate, and carrying capacity as functions of time and distance and provide parameter selection recommendations; 2) develop new distance metrics such as shared interest between users [27]; 3) categorizing news stories into clusters and modeling each cluster; 4) modeling more complex scenarios such as a news with multiple sources; 5) evaluate the proposed model with other social media such as Twitter.

ACKNOWLEDGMENT

We would like to thank Kristina Lerman for making the Digg 2009 data set available to our research project. This work was supported in part by Arizona State University New College SRCA grants.

REFERENCES

[1] A. L. Hughes and L. Palen, "Twitter Adoption and Use in Crisis Twitter Adoption and Use in Mass Convergence and Emergency Events," in *Proceedings of the 6th International Information Systems for Crisis Response and Management Conference*, 2009.

[2] K. Lerman, and R. Ghosh, "Information Contagion: an Empirical Study of Spread of News on Digg and Twitter Social Networks," in *Proceedings of International Conference on Weblogs and Social Media (ICWSM)*, 2010.

[3] G. V. Steeg, R. Ghosh, and K. Lerman, "What Stops Social Epidemics?" in *Proceedings of International AAAI Conference on Weblogs and Social Media*, 2011.

[4] S. Ye and F. Wu, "Measuring Message Propagation and Social Influence on Twitter.com," 2010.

[5] M. Cha, A. Mislove, and K. P. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network," in *Proceedings of the 18th international conference on World wide web*, ser. WWW '09, 2009.

[6] J. Yang and J. Leskovec, “Modeling Information Diffusion in Implicit Networks,” in *Proceedings of IEEE International Conference on Data Mining*, 2010.

[7] R. Ghosh and K. Lerman, “A Framework for Quantitative Analysis of Cascades on Networks,” in *Proceedings of Web Search and Data Mining Conference (WSDM)*, 2011.

[8] M. Gomez Rodriguez, J. Leskovec, and A. Krause, “Inferring networks of diffusion and influence,” in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2010.

[9] J. D. Murray, *Mathematical Biology I. An Introduction*. Springer-Verlag, 1989.

[10] O. Diekmann and J. A. P. Heesterbeek, “Mathematical epidemiology of infectious diseases: Model building, analysis and interpretation,” *Wiley Series in Mathematical & Computational Biology*, 2000.

[11] C. Gerald, and P. Wheatley, *Applied Numerical Analysis*. Addison-Wesley, 1994.

[12] K. Lerman, “Digg 2009 data set,” http://www.isi.edu/~lerman/downloads/digg2009.html.

[13] D. Agrawal, C. Budak, and A. E. Abbadi, “Tutorial: Information diffusion in social networks: Observing and influencing societal interests,” 2011.

[14] S. Tang, N. Blenn, C. Doerr, and P. V. Mieghem, “Digging in the Digg Social News Website,” *IEEE Transactions on Multimedia*, 2011.

[15] D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins, “Information diffusion through blogspace,” in *Proceedings of the 13th international conference on World Wide Web*, ser. WWW ’04, 2004.

[16] M. Cha, A. Mislove, B. Adams, K. Gummadi, “Characterizing Social Cascades in Flickr,” in *Proceedings of ACM SIGCOMM Workshop on Social Networks (WOSN)*, 2008.

[17] B. Yu, and H. Fei, “Modeling Social Cascade in the Flickr Social Network,” in *Proceedings of International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 2009.

[18] K. Saito, M. Kimura, K. Ohara, and H. Motoda, “Efficient Discovery of Influential Nodes for SIS Models in Social Networks,” *Knowledge and Information Systems*, 2011.

[19] D. Kempe, J. Kleinberg, and E. Tardos, “Maximizing the Spread of Influence through a Social Network,” in *Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2003.

[20] M. Gomez Rodriguez, J. Leskovec, and A. Krause, “Inferring networks of diffusion and influence,” in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, ser. KDD ’10, 2010.

[21] C. Budak, D. Agrawal, and A. El Abbadi, “Limiting the spread of misinformation in social networks,” in *Proceedings of the 20th international conference on World wide web*, ser. WWW ’11, 2011.

[22] J. Yang, and S. Counts, “Comparing Information Diffusion Structure in Weblogs and Microblogs,” in *Proceedings of International AAAI Conference on Weblogs and Social Media*, 2010.

[23] J. Iribarren, and E. Moro, “Impact of Human Activity Patterns on the Dynamics of Information Diffusion,” *Physical Review Letters*, 2009.

[24] Z. Katona, P. Zubcsek, and M. Sarvary, “Network Effects and Personal Influences: Diffusion of an Online Social Network,” *Journal of Marketing Research*, 2011.

[25] R. Cantrell and C. Cosner, *Spatial Ecology via Reaction-Diffusion Equation*. Wiley, 2003.

[26] P. Fife, “Spatial Ecology via Reaction-Diffusion Equation,” *Mathematical Aspects of Reacting and Diffusing Systems*, 1979.

[27] F. Wang, K. Xu, and H. Wang, “Discovering shared interests in online social networks,” *Proceedings of IEEE ICDCS Workshop on Peer-to-Peer Computing and Online Social Networking (HOTPOST)*, 2012.