2017 IEEE 37th International Conference on Distributed Computing Systems Workshops

Modeling the Spread of Influence for Independent Cascade Diffusion Process in Social Networks

Zesheng Chen and Kurtis Taylor
Department of Computer Science
Indiana University - Purdue University Fort Wayne, Indiana 46805
Email: chenz@ipfw.edu and taylkj04@students.ipfw.edu

***Abstract*—Modeling the spread of influence in online social networks is important for predicting the influence of individuals and better understanding many scenarios in social networks, such as the influence maximization problem. The previous work on modeling the spread of influence makes the assumption that the statuses of nodes in a network are independent of each other, which is apparently not correct for social networks. The goal of this work is to derive an accurate mathematical model to characterize the spread of influence for the independent cascade diffusion process in online social networks. Specifically, we apply the susceptible-infected-recovered epidemic model from epidemiology to characterize the independent cascade diffusion process and derive a general mathematical framework. To approximate the complex spatial dependence among nodes in a network, we propose a Markov model to predict the spread of influence. Through the extensive simulation study over several generated topologies and a real coauthorship network, we show that our designed Markov model has much better performance than the existing independent model in predicting the influence of individuals in online social networks.**

***Index Terms*—Social influence, online social networks, independent cascade diffusion, spatial dependence, Markov dependence.**

I. Introduction

Influence can spread through an online social network. "Word-of-mouth" and "viral marketing" effects have been widely exploited to promote new products and technological innovations. For example, when an individual adopts a new product and finds it useful, she or he may post the information to her or his Facebook and recommend it to friends and colleagues. One of this individual's friends takes the advice, and may also feel excited about the product and spread the words about it to her or his own friends through Facebook. In such a way, social influence can help diffuse new products or ideas. A classic research problem on the spread of influence is called the *influence maximization* problem, which studies how to choose few key individuals in a social network to give free samples of a product so that it can maximize the number of the individuals who will eventually adopt the product [1], [2], [3], [4], [5]. There are two basic influence diffusion processes that have been widely investigated: independent cascade and linear threshold [1]. In this paper, we will focus on the independent cascade diffusion process.

The spread of influence in online social networks bears resemblance to epidemic processes in networks, such as malware spread and information dissemination [6], [7], [8], [9], [10], [11], [12]. Thus, epidemic models have been applied to modeling the influence propagation in social networks [13], [14], [15]. The mathematical models in previous work, however, make the assumption that the statuses of nodes in a network are independent of each other. Such an assumption is apparently *not* accurate. For example, two friends in an online social network tend to both either adopt a product or reject it. That is, the statuses of nodes in online social networks are spatially positively correlated.

The goal of this work is to find an accurate mathematical model to characterize the spread of influence for the independent cascade diffusion process in online social networks. Specifically, we attempt to apply a mathematical model to answer the question: given an online social network and an individual (or several individuals) who initially adopt(s) a product, what is the expected number of people who will eventually adopt the product? The answer to this question can potentially help better understand many scenarios on social influence, such as the influence maximization problem. Specifically, the traditional way for the influence maximization strategy to predict the influence of initially selected nodes is to use simulations [1], [3], [5], which usually take a long time to run. Our proposed mathematical model can potentially provide the same or similar prediction in a significantly shorter time.

In this work, we first apply the susceptible-infected-recovered model from epidemiology to characterize the independent cascade diffusion process. We then focus on approximating the complex spatial dependence among nodes in a network. Specifically, we apply a spatial Markov dependence assumption and propose a new mathematical model. Finally, we use simulations to evaluate the performance of our model. We summarize our contributions in the following:

- We derive a general mathematical framework to characterize the spread of influence in online social networks and point out the difficulty in finding an accurate model.
- We consider two different approximations for the above model: the independent model and the Markov model. The independent model is based on the spatial independence assumption that has been applied to the previous work [14], [15], whereas the Markov model is inspired by the Markov Random Field and considers a certain spatial dependence. Through extensive simulations over several generated topologies (*i.e.*, a power-law topology, a lattice, an ER random graph, and an exponential growth

1545-0678/17 $31.00 © 2017 IEEE
DOI 10.1109/ICDCSW.2017.28

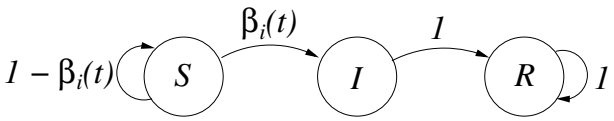


Fig. 1. SIR model for node i.

random graph) and a real coauthorship network, we show that the Markov model better characterizes the spread of influence than the independent model and can accurately predict the number of people who will eventually adopt the product.

The remainder of this paper is structured as follows. Section II introduces the independent cascade diffusion process. Section III derives a general mathematical framework and proposes a Markov model, and Section IV evaluates the performance of our designed Markov model and compares it with the independent model and simulation results. Finally, Section V presents the related work and discussions, and Section VI concludes this paper.

II. Independent Cascade Diffusion Process

The independent cascade (IC) diffusion process has been widely used to describe the procedure of influence diffusion in social networks [1]. Specifically, consider a discrete-time system. A set of nodes are initially infected (or active). If a node i first becomes infected at time t, it will have a single chance to infect its susceptible neighbor j (that is not infected yet) with an influence probability β_{ij}. If node i succeeds, then node j will become infected at time $t+1$. After time t, node i cannot make any further attempts to infect its neighbors. Such a diffusion process continues until no more infections are possible [1].

The IC diffusion process can be well characterized by a *susceptible-infected-recovered* (SIR) mathematical model from epidemiology [13]. As shown in Figure 1, node i has three statuses: susceptible, infected, and recovered. At time t, if node i is susceptible, it can either be infected by its infectious neighbors with probability $\beta_i(t)$ or stay in the susceptible status with probability $1-\beta_i(t)$. If node i is infected at time t, it can attempt to infect its susceptible neighbors, but it will become recovered at time $t+1$ with probability 1. If node i is recovered, it cannot infect any other nodes and will stay in the recovered status forever.

III. Mathematical Models

In this section, we first discuss the system model and topologies. Next, we present a general mathematical framework to describe the spread of influence in online social networks using the IC diffusion process and point out the difficulty in finding the exact solution. Finally, we consider two different spatial approximation methods to derive the recursive equations: the independent model and the Markov model.

A. System Model

Let $G(V,E)$ represent an online social network, where V is the set of nodes and E is the set of links. We consider directed graphs and treat undirected graphs as special cases of directed graphs. In an undirected graph, if a link $(i,j) \in E$, then $(j,i) \in E$. Let $N_i = \{j|(j,i) \in E\}$ denote a neighborhood of node i, *i.e.*, a set of nodes that can influence node i directly. Let $n = |V|$ denote the number of nodes in the network. Similarly, $|N_i|$ denotes the number of neighbors of node i.

Different networks may have different topological structures. Typical generated topologies include lattice, ER random, exponential growth random, and power-law graphs. A lattice graph is a regular topology, in which each node has four neighbors [9]. An ER random graph is formed by considering all possible pairs of nodes and connecting a pair with a probability [16]. An exponential growth random graph is built through a sequential growth model that adds nodes one by one to the graph and connects the newly added node to each existing node with equal likelihood (*i.e.*, uniform attachment) [17], [18]. A power-law topology is similar to the exponential growth random graph, but the probability that the newly added node connects to an existing node is based on the nodal degree of this existing node (*i.e.*, preferential attachment) [17], [19].

B. Mathematical Framework

We apply the SIR model to characterize the spread of influence in online social networks. Specifically, let $X_i(t)$ denote the status of node i at time t, *i.e.,*

$$X_i(t) = \begin{cases} 0, & \text{if node } i \text{ is susceptible at time } t; \\ 1, & \text{if node } i \text{ is infected at time } t; \\ -1, & \text{if node } i \text{ is recovered at time } t. \end{cases} \quad (1)$$

Similarly, let $\mathbf{X}_{N_i}(t)$ denote the statuses of node i's neighbors at time t. There are $3^{|N_i|}$ possible combinations for $\mathbf{X}_{N_i}(t)$. Let $S_i(t)$, $I_i(t)$, and $R_i(t)$ denote the probabilities that node i is susceptible, infected, and recovered at time t ($t \geq 0$), where $i = 1, 2, \cdots, n$. Note that $S_i(t)+I_i(t)+R_i(t) = 1$. The status of node i at time $t+1$ can be expressed based on its status and its neighbors' statuses at time t (as shown in Figure 1), *i.e.*,

$$S_i(t+1) = S_i(t)[1-\beta_i(t)] \quad (2)$$

$$I_i(t+1) = S_i(t)\beta_i(t) \quad (3)$$

$$R_i(t+1) = I_i(t) + R_i(t) \quad (4)$$

where $\beta_i(t)$ is the probability that node i will be infected by its infectious neighbors at time $t+1$, given that it is susceptible at time t, *i.e.*,

$$\begin{aligned} &\beta_i(t) \\ &= P(X_i(t+1) = 1 | X_i(t) = 0) \\ &= \sum_{\mathbf{x}_{N_i}(t)} P(X_i(t+1) = 1, \mathbf{X}_{N_i}(t) = \mathbf{x}_{N_i}(t) | X_i(t) = 0) \\ &= \sum_{\mathbf{x}_{N_i}(t)} P(\mathbf{X}_{N_i}(t) = \mathbf{x}_{N_i}(t) | X_i(t) = 0) \cdot f_i(t) \end{aligned} \quad (5)$$

where

$$f_i(t) = P(X_i(t+1) = 1 | \mathbf{X}_{N_i}(t) = \mathbf{x}_{N_i}(t), X_i(t) = 0), \quad (6)$$

which represents the probability that the node i's neighbors can infect the susceptible node i and depends on the diffusion process.

If the IC diffusion process is considered, node i can be infected only by its neighbors that are in the infected status, *i.e.*, $x_j(t) = 1$, $j \in N_i$. If neighbor j of node i is susceptible or recovered (*i.e.*, $x_j(t) = 0$ or -1), the probability that node j will *not* infect node i is 1 (*i.e.*, $(1-\beta_{ji})^0$); otherwise, neighbor j is infected (*i.e.*, $x_j(t) = 1$), and it will *not* infect node i with the probability $1-\beta_{ji}$ (*i.e.*, $(1-\beta_{ji})^1$). Combining these two cases together, we can find that the probability that neighbor j will *not* infect node i at time t is $(1-\beta_{ji})^{\frac{x_j^2(t)+x_j(t)}{2}}$, where $x_j(t) = -1$, 0, or 1. Since all infected neighbors of node i attempt to infect node i independently, the probability that node i will *not* be infected by its neighbors at time t becomes $\prod_{j\in N_i}(1-\beta_{ji})^{\frac{x_j^2(t)+x_j(t)}{2}}$, which leads to

$$f_i^{IC}(t) = 1 - \prod_{j\in N_i} (1-\beta_{ji})^{\frac{x_j^2(t)+x_j(t)}{2}} \quad (7)$$

representing the probability that the node i's neighbors can infect susceptible node i for the IC diffusion process.

The difficulty in finding $\beta_i(t)$ lies in computing the conditional joint distribution $P(\mathbf{X}_{N_i}(t) = \mathbf{x}_{N_i}(t) | X_i(t) = 0)$, *i.e.*, given node i is susceptible at time t, computing the joint distribution of all its neighbors' statuses at time t. If node i has k neighbors, the number of the possible combinations of these statuses is 3^k, which is too expensive to compute. To reduce the computation, we consider two approximations to this conditional joint distribution.

C. Independent Model

The simplest way to approximate the distribution is to assume that all nodes are *independent* at time t, which has been applied in previous work [14], [15]. That is, assume that the statuses of node i and its neighbors are independent, which leads to

$$P(\mathbf{X}_{N_i}(t) = \mathbf{x}_{N_i}(t) | X_i(t) = 0) = \prod_{j\in N_i} P(X_j(t) = x_j(t)). \quad (8)$$

Putting this equation and Equation (7) into Equation (5), we can find the simplified $\beta_i(t)$ for the IC diffusion process:

$$\begin{aligned}
&\beta_i^{IC_ind}(t) \\
&= \sum_{\mathbf{x}_{N_i}(t)} \prod_{j\in N_i} P(X_j(t) = x_j(t)) \cdot f_i^{IC}(t) \\
&= 1 - \prod_{j\in N_i} \sum_{x_j(t)} P(X_j(t) = x_j(t))(1-\beta_{ji})^{\frac{x_j^2(t)+x_j(t)}{2}} \\
&= 1 - \prod_{j\in N_i} [1 - \beta_{ji} P(X_j(t) = 1)] \\
&= 1 - \prod_{j\in N_i} (1 - \beta_{ji} I_j(t)) . \quad (9)
\end{aligned}$$

Putting $\beta_i^{IC\_ind}(t)$ into Equations (2) and (3), we can find $S_i(t)$, $I_i(t)$, and $R_i(t)$ recursively from $S_i(0)$, $I_i(0)$, and $R_i(0)$.

Comparing with the previous work, we find that the unified model proposed in [14] for the IC diffusion process (*i.e.*, Equation (11) in [14]) is the *same* as the independent model in Equation (9). Moreover, the inclusion-exclusion theorem shown in [15] (*i.e.*, Theorem 1 and Equation (9) in [15]) is *indeed* the independent model by expanding the product term in Equation (9). Therefore, the independent model is essentially the model considered in the previous work [14], [15]. Although the independent model has been proposed in the previous work, in this paper we derive it under a general mathematical framework.

D. Markov Model

In the Markov model, we assume that given node i is susceptible, the statuses of its neighbors are *conditionally* independent, *i.e.*,

$$\begin{aligned}
&P(\mathbf{X}_{N_i}(t) = \mathbf{x}_{N_i}(t) | X_i(t) = 0) \\
&= \prod_{j\in N_i} P(X_j(t) = x_j(t) | X_i(t) = 0). \quad (10)
\end{aligned}$$

Such an approximation is inspired by the local Markov property in Markov Random Field and has been applied in modeling malware propagation in networks [10]. The intuition behind Equation (10) is that the status of node i is *not* independent of its neighbors, but is indeed affected by them. As pointed out in the Introduction, two friends in a social network tend to agree upon the decision of adopting or rejecting a product. The spatial dependence of two friends is positively correlated. Moreover, since the status of neighbor j of node i depends on its own neighbors, the status of node i can rely on the neighbors of node j. In such a way, all nodes in a connected graph can depend on each other spatially. To avoid the extensive computation on the complex spatial dependence, in this paper we consider the one-step spatial dependence as shown in Equation (10), *i.e.*, spatial conditional independence.

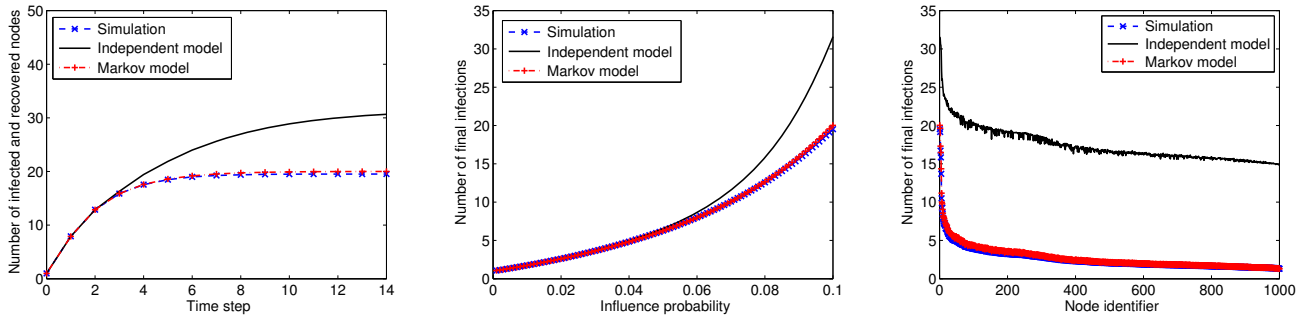
Considering the IC diffusion process,, we have

$$\begin{aligned}
&\beta_i^{IC_mar}(t) \\
&= \sum_{\mathbf{x}_{N_i}(t)} \prod_{j\in N_i} P(X_j(t) = x_j(t) | X_i(t) = 0) \cdot f_i^{IC}(t) \\
&= 1 - \prod_{j\in N_i} [1 - \beta_{ji} P(X_j(t) = 1 | X_i(t) = 0)] . \quad (11)
\end{aligned}$$

To find $\beta_i^{IC\_mar}(t)$, we need to derive the expression of $P(X_j(t) = 1 | X_i(t) = 0)$. Set $M(i,j,t) = P(X_i(t) = 0, X_j(t) = 0)$ and $N(i,j,t) = P(X_i(t) = 0, X_j(t) = 1)$. Moreover, use the notation

$$\beta_{i/j}(t) = \sum_{\mathbf{x}_{N_i'}(t)} P(\mathbf{X}_{N_i'}(t) = \mathbf{x}_{N_i'}(t) | X_i(t) = 0) \cdot f_i(t) \quad (12)$$

where $N_i' = N_i - \{j\}$. That is, consider the effect of neighbors without node j.



(a) Transient behavior ($\beta = 0.1$) (b) Influence of the node with the largest nodal degree (c) Nodes' influence in descending order ($\beta = 0.1$)

Fig. 2. Influence spreads over a BA power-law topology ($n = 1,000$).

Note that

$$\begin{aligned} &M(i,j,t) \\ &= P(X_i(t)=0, X_j(t)=0, X_i(t-1)=0, X_j(t-1)=0) \\ &= M(i,j,t-1)(1-\beta_{i/j}(t-1))(1-\beta_{j/i}(t-1)) \end{aligned} \quad (13)$$

and

$$\begin{aligned} &N(i,j,t) \\ &= P(X_i(t)=0, X_j(t)=1, X_i(t-1)=0, X_j(t-1)=0) \\ &= M(i,j,t-1)(1-\beta_{i/j}(t-1))\beta_{j/i}(t-1). \end{aligned} \quad (14)$$

Thus,

$$P(X_j(t)=1|X_i(t)=0) = \frac{N(i,j,t)}{S_i(t)}. \quad (15)$$

Assuming the initial condition (*i.e.*, $S_i(0)$, $I_i(0)$, $R_i(0)$, $M(i,j,0)$, and $N(i,j,0)$), we can find $S_i(t)$, $I_i(t)$, $R_i(t)$, $M(i,j,t)$, and $N(i,j,t)$ recursively.

IV. Simulation Results and Performance Evaluation

In this section, we evaluate the performance of two models on predicting the spread of influence in online social networks through simulations. Specifically, we first discuss the simulation setup. We then evaluate the performance of models in both generated topologies and a real topology.

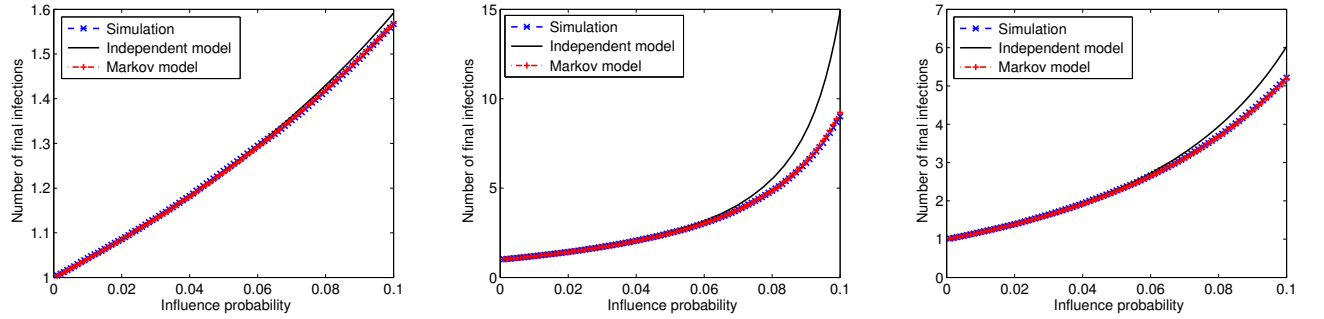
A. Simulation Setup

We simulate the spread of influence for the IC diffusion process in an undirected graph and assume that the influence probability is the same for all links (*i.e.*, $\beta_{ij} = \beta, \forall (i,j) \in E$). A discrete-time system is considered. In each time step, if a node is susceptible, one of its infected neighbors can infect it with an influence probability β. If this node is infected, it can potentially infect its neighbors and will recover at the next time step. Otherwise, the node is recovered; and it cannot infect any other nodes and will stay in the recovered status forever. We run 20,000 times using different seeds for each scenario and average the number of final infections (*i.e.*, number of recovered nodes at the end) over these 20,000 runs. In each run, the simulation starts from a single infected node and continues running until no more infections are possible. If a topology is not connected, we consider the spread of influence in the giant component. For the influence probability β, similar to [2], we assume $0.001 \leq \beta \leq 0.1$. Specifically, the case of $\beta = 0.1$ indicates high influence.

B. Generated Topologies

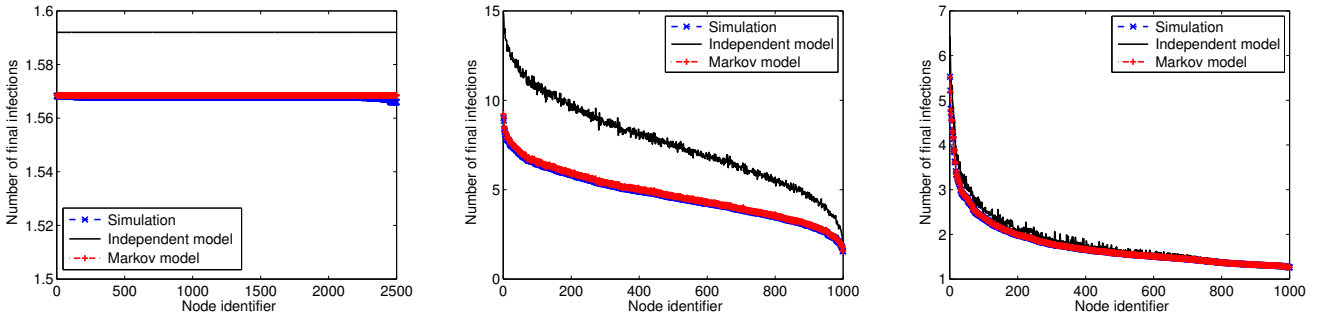
We compare the performance of mathematical models (*i.e.*, independent model and our proposed Markov model) with simulation results for several generated topologies. We first consider a Barabási-Albert (BA) power-law topology generated by the BRITE tool [20]. The topology contains 1,000 nodes and has an average nodal degree of 3.99. Figure 2(a) shows how the number of infected and recovered nodes varies with time, when the infection starts from a single node with the largest nodal degree and the influence probability is 0.1 (*i.e.*, $\beta = 0.1$). It can be seen that the independent model overestimates the spread ability of the infection, whereas the Markov model accurately predicts the spread process. Figure 2(b) indicates how the number of final infections changes with the influence probability β ($0.001 \leq \beta \leq 0.1$). Similarly, the initially infected node is the one with the largest nodal degree. It can be seen that when β is small, both models work well. When β is large (*e.g.*, $\beta > 0.06$), however, the performance of the Markov model is much better than that of the independent model. Figure 2(c) shows the number of final infections for every node selected as the starting node, when $\beta = 0.1$. The number is in descending order, according to the simulation results. Obviously, the Markov model can accurately estimate the number of final infections when the starting node is an arbitrary node.

Next, we simulate the spread of influence in other generated topologies. Figures 3 and 4 shows the influence propagation in the lattice, ER random, and exponential growth random graphs. In the simulation, the lattice graph contains 2,500 nodes, each of which has four neighbors. The ER random graph contains 1,000 nodes and has an average nodal degree of 7.91. The exponential growth random graph is with 1,000 nodes and an average nodal degree of 3.99. Similar to Figure 2(b), Figure 3 indicates how the number of final



(a) Lattice ($n = 2,500$) (b) ER random graph ($n = 1,000$) (c) Exponential growth random graph ($n = 1,000$)

Fig. 3. Influence of the node with the largest nodal degree in different generated topologies.



(a) Lattice ($n = 2,500$) (b) ER random graph ($n = 1,000$) (c) Exponential growth random graph ($n = 1,000$)

Fig. 4. Nodes' influence in descending order in different generated topologies ($\beta = 0.1$).

infections varies with the influence probability β, when the initially infected node is the one with the largest nodal degree. Moreover, similar to Figure 2(c), Figure 4 plots the number of final infections for every node selected as the starting node, when $\beta = 0.1$. It can be seen that in all generated topologies, the independent model overestimates the influence ability of a node, whereas the results from the Markov model overlap with those from simulations. Therefore, the Markov model performs much better than the independent model. It is noted that due to the symmetric property of a regular graph, the number of final infections are (almost) identical for all nodes in the lattice topology, as shown in Figure 4(a).

C. A Real Topology

We further consider a real topology that is a coauthorship network of scientists working on network theory and experiment [21]. Since this real topology is not connected, we only consider the giant component with 379 nodes and an average nodal degree of 4.82. Figure 5 plots how the number of final infections changes with the influence probability β when the initially infected node is the one with the largest nodal degree, whereas Figure 6 shows the number of final infections for every node selected as the starting node when $\beta = 0.1$. Similar to Figures 2, 3, and 4, we can see that the Markov model closely follows the simulation results and better characterizes the spread of influence in the coauthorship network than the independent model.

V. Related Work and Discussions

The spread of information or malware in networks has been studied in [9], [10], [11], [12]. The epidemiological model considered in these works is usually either the susceptible-infected (SI) model or the susceptible-infected-susceptible (SIS) model. As pointed out in [13], the IC diffusion process in social networks is better characterized by the SIR model.

The spread of influence for the IC diffusion process in social networks has been modeled in [14], [15]. In the previous work, however, it is assumed that all nodes in a network are independent of each other. In this paper, we have shown that the models in [14], [15] are essentially the independent model under the general mathematical framework.

One application of modeling the spread of influence in social networks is for the influence maximization problem (IMP), which has been widely studied [1], [2], [3], [4], [5]. The conventional approach to evaluate a strategy for IMP is to apply simulations. For instance, in [1], each of different strategies (*i.e.*, greedy, high degree, central, and random) runs 10,000 times in simulations, which take a long time to get the results. An accurate mathematical model can shorten the time significantly. For example, in our performance evaluation over

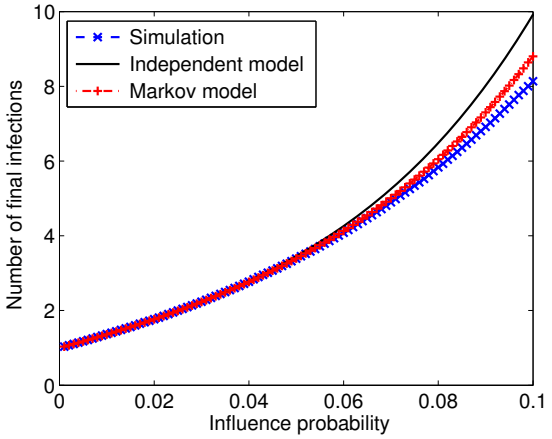


Fig. 5. Influence of the node with the largest nodal degree in a coauthorship network ($n = 379$).

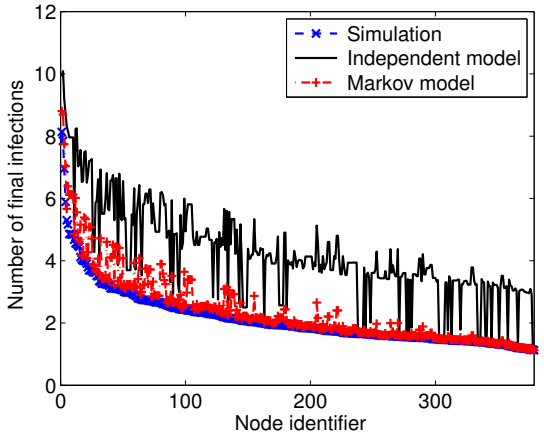


Fig. 6. Nodes' influence in descending order in a coauthorship network ($n = 379$ and $\beta = 0.1$).

the coauthorship network using a computer with high power processors, the simulation took about 374 seconds to get the average of 20,000 runs, whereas the Markov model used only 6 seconds. Therefore, our proposed model can complement the solutions to IMP.

VI. Conclusions

In this work, we have explored the spatial Markov dependence among nodes and proposed a new mathematical model to describe the spread of influence for the IC diffusion process in online social networks. Our designed model takes into consideration a certain spatial dependence. Through extensive simulations over several generated topologies and a real coauthorship network, we have shown that our proposed Markov model better characterizes the process of IC diffusion in online social networks than the existing independent model.

As our on-going work, we plan to extend our model for the linear threshold diffusion process described in [1].

ACKNOWLEDGMENTS

This work was supported in part by a 2016 Summer Faculty Grant from the Purdue Research Foundation and a 2016 IPFW Undergraduate Summer Research Support Grant.

References

[1] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *Proc. of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2003, Washington, DC, pp. 137-146.

[2] C. Wang, W. Chen, and Y. Wang, "Scalable influence maximization for independent cascade model in large-scale social Networks," in *Data Mining and Knowledge Discovery*, Springer, May 2012.

[3] W. Chen, L. V.S. Lakshmanan, and C. Castillo, "Information and influence propagation in social networks," in *Morgan & Claypool Synthesis Lectures on Data Management*, Morgan & Claypool, Oct. 2013.

[4] X. He and D. Kempe, "Robust influence maximization," in *Proc. of the 22nd SIGKDD Conference on Knowledge Discovery and Data Mining*, Aug. 2016, San Francisco, CA.

[5] G. Tong, W. Wu, S. Tang, and D. Du, "Adaptive influence maximization in dynamic social networks," *IEEE/ACM Transactions on Networking*, vol. 25, no. 1, Feb. 2017, pp. 112-125.

[6] D. J. Daley and J. Gani, *Epidemic Modeling: An Introduction*. Cambridge University Press, 1999.

[7] R. Pastor-Satorras, C. Castellano, P. V. Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Rev. Mod. Phys.*, vol. 87, no. 3, Aug. 2015, pp. 925.

[8] Y. Wang, S. Wen, Y. Xiang, and W. Zhou, "Modeling the propagation of worms in networks: A survey", *IEEE Communications Surveys & Tutorials*, vol. 16, no. 2, Second Quarter 2014, pp. 942-960.

[9] J. O. Kephart and S. R. White, "Directed-graph epidemiological models of computer viruses," in *Proc. of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy*, 1991, pp. 343-359.

[10] Z. Chen and C. Ji, "Spatial-temporal modeling of malware propagation in networks," *IEEE Transactions on Neural Networks: Special Issue on Adaptive Learning Systems in Communication Networks*, vol. 16, no. 5, Sept. 2005.

[11] S. M. Cheng, W. C. Ao, P. Y. Chen, and K. C. Chen, "On modeling malware propagation in generalized social networks," *IEEE Communications Letters*, vol. 15, no. 1, Jan. 2011, pp. 25-27.

[12] Y. F. Chou, H. H. Huang, and R. G. Cheng, "Modeling information dissemination in generalized social networks," *IEEE Communications Letters*, vol. 17, no. 7, July. 2013, pp. 1356-1359.

[13] B. A. Prakash, D. Chakrabarti, M. Faloutsos, N. Valler, and C. Faloutsos, "Threshold conditions for arbitrary cascade models on arbitrary networks," in *Proc. of the 2011 IEEE 11th International Conference on Data Mining*, Vancouver, pp. 537-546.

[14] A. Srivastava, C. Chelmis, and V. K. Prasanna, "Influence in social networks: A unified model?" in *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Aug. 2014, Beijing, China, pp. 451 - 454.

[15] M. Zhang, C. Dai, C. Ding, and E. Chen, "Probabilistic solutions of influence propagation on social networks," in *Proc. of the 22nd ACM international conference on Information & Knowledge Management*, Oct. 2013, San Francisco, CA, pp. 429-438.

[16] P. Erdös and A. Rényi, "On the evolution of random graphs," in *Publ. Math. Inst. Hung. Acad. Sci.*, 1960, vol. 5, pp. 17-61.

[17] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, Oct. 1999, pp. 509-512.

[18] M. O. Jackson and B. W. Rogers, "Meeting strangers and friends of friends: How random are socially generated networks?" *American Economic Review*, vol. 97, no. 3, June 2007, pp 890-915.

[19] A.-L. Barabási, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks," *Physica A*, vol. 272, 1999, pp. 173-187.

[20] BRITE, Power-Law Topology Generator [Online]. Available: http://www.cs.bu.edu/brite/ (January/2017 accessed).

[21] Collaboration Network in Science of Networks [Online]. Available: http://vlado.fmf.uni-lj.si/pub/networks/data/collab/netscience.htm (January/2017 accessed).