Academic Social Network Information Dissemination Model

Huaqiang Jiang[1, a], Yuncui Pan[2,b] and Xiang Jie[1,c] and Jianmei Miao[3,d]
[1]College of Info Science and Engineering of Hangzhou Normal University, China
[2]School of Foreign Languages, Anhui Science and Technology University, China
[3] Science and Research Department of Hangzhou Normal University, China
[a]jhq@hznu.edu.cn, [b]cindypanpan@163.com, [c]xia_ngj_ie@hotmail.com, [d]mjm@hznu.edu.cn

***Abstract*—This paper constructs an academic social network information dissemination model. A social network dataset was established through acquiring the user relation in the website which named ResearchGate. The diagram of information dissemination trend were drew according to a dissemination statistical of 30 questions within 30 days, and through a calculation of the probability of information dissemination among the network, the regularity of information dissemination in the network was obtained.**

Keywords-Academic social networks, Information dissemination characteristics, Information dissemination model

I. INTRODUCTION

As the rapid increase of the number of literature and fast development of social network website, academic social networks (ASN), a kind of web-based dissemination platform, was established to enhance academic information dissemination and improve study efficiency [1]. ASN makes use of internet to provide academic study assisted tools and service to pool researchers' dissemination needs during academic study, and its main audience are teachers, students, researchers, etc. ASN not only makes researchers easy to contact with each other and improve the dissemination efficiency, but also helps to improve the information acquisition speed and accuracy through academic recommendation, and better promote the spread and impact of the academic achievements.

II. INFORMATION DISSEMINATION MODEL

ASN is user-oriented rather than content-oriented, which gets rid of traditional academic information dissemination model. User-oriented model relies on relationships among people and contributes to disseminate academic information. Study of ASN information dissemination model helps to reach a better understanding of the information dissemination process and behavior, of the analysis of academic research trend, innovation of academic achievement evaluation methods, etc.

A. ASN information dissemination process

The development of social network makes changes on academic information production, publication, dissemination, evaluation and other aspects. ASN is constituted by users, contents, relationships and tools, and makes researchers deliver research information easily and simply. Subscription mechanism makes variety relationships operation and snowballing feedback become possible. Take ResearchGate, one of the largest science social network website, as an example (http://www.researchgate.net/)[2], it contains functions like: Live Feed, Topics, Publications, Projects, etc. As to the process of dissemination, it can be divided into concerning process, questioning process, reviewing process and tracking process. In the concerning process, user u can draw user f's interest through personal profile, system recommendation (generally from colleagues in the same institution) or raise questions in Topics, and then user f become user u's follower (fans). This process can be realized without user u's consent, afterwards, all of user u's updated information will be synchronized appeared in user f's Live Feed. In the questioning process, user can ask questions or post messages (limited to 150 words, but can be annotated text, papers, hyperlinks, etc.) anywhere and anytime in Topics. In the reviewing process, user can comment on any question (or message) and also can be annotated text, papers, hyperlinks, etc. In the tracking process, user can track interested questions with concerning function, and all the being tracked information will be updated to the Live Feed page.

Therefore, the dissemination process of messages on the ASN can be summarized as follows:

(1) User u post messages on Topics;

(2) User u's followers acquire messages through content alert or actively check user u's page on Live Feed;

(3) User u's other followers can reply, review, track on user f, or use shared information to spread the messages again;

(4) Repeat the process (2) and (3).

B. Academic social network information dissemination model description

According to the analysis of the regularity of ASN information dissemination, we define each user as a node and the concerning relationship between each node can be defined as the edges, and information only goes along the edges. One message delivered by a user can be seen, shared, or disseminated by his followers; however, followers who have no interest in the messages will become "immune" and messages will no longer be spread. In this regard, we divide the node in the network into three

Published by Atlantis Press, Paris, France.

categories: dissemination node, uninfected node, immune node. Dissemination node is the one receive neighbor node's messages and capable of spreading them, that's to say it has probability of spreading; uninfected node is the one don't have received messages from neighbor node and is likely to accept the messages, that is it has probability of being infected; immune node is the one has received neighbor node's message, but don't have the ability to disseminate.

Based on the above-mentioned analysis, to establish an ASN information dissemination model [3] and describe the changing process of messages with time, define the following basic conditions is necessary:

(1) Define the total number of nodes as *N*, the proportion of uninfected node, dissemination node, and immune node in the total node *N*, respectively, are *s(t)* ,*i(t)* and *r(t)*, *t* is dissemination time.

(2) After a dissemination node contact with an uninfected node, the uninfected node will become a dissemination node with the probability λ.

(3) After a dissemination node contact with an immune node, the dissemination node will become an immune node with the probability μ.

According to the above-mentioned, obviously there is $i(t) + s(t) + r(t) = 1$.Newman et al[4,5]. have proven interpersonal networks are small-world networks[4,5], and degree of node is close to the average degree $\bar{k}$. Making a reference to infectious disease dynamics model [6], there is the following information dissemination model equationsEq.1 –Eq.3:

$$N\frac{ds}{dt} = -\lambda \bar{k} Nsi \quad (1)$$

$$N\frac{di}{dt} = \lambda \bar{k} Nsi - \mu Ni \quad (2)$$

$$N\frac{dr}{dt} = \mu Ni \quad (3)$$

Suppose the proportion of uninfected node at the initial time is $s0\ (s0 > 0)$, dissemination node is $i0\ (i0 > 0)$, immune node is $r0 = 0$, then initial value is Eq.4:

$$\begin{cases} \frac{ds}{dt} = -\lambda \bar{k} si, & s(0) = s_0 \\ \frac{di}{dt} = \lambda \bar{k} si - \mu i, & i(0) = i_0 \\ \frac{dr}{dt} = \mu i, & r(0) = 0 \end{cases} \quad (4)$$

From $s + i + r = 1$ can get $dr/dt = -di/dt - ds/dt$, then model equation can simplified as Eq.5 :

$$\begin{cases} \frac{ds}{dt} = -\lambda \bar{k} si, & s(0) = s_0 \\ \frac{di}{dt} = \lambda \bar{k} si - \mu i, & i(0) = i_0 \end{cases} \quad (5)$$

The model equation can not get analytic solution, *s(t)* , *i(t)* and *r(t)'s* numerical solution can be obtained through numerical simulation method.

III. MODEL VALIDATION AND ANALYSIS

In order to verify the information dissemination model established in this paper, we conduct research into the regularity of information dissemination and take ReserchGate as subject.

A. Experimental data acquisition

First, ASN user relationship diagram establishment. We use Python language to get a web crawler script page, use a registered user's identity to grasp page, and use snowball sampling method to implement breadth-first search (BFS) algorithm to extract users' information[7]. Starting from the specific user to obtain user's followers and list of concerning users, then added to the queue; further explore the followers or concerning users to obtain a list of followers and concerning users, then added to the queue; until reach the desired number of users. The generated network is one of the sub-network of ResearchGate, and it has a digraph denoted as G=(V ,E), node $v \in V$ indicates a user's ID, edge $(u, f) \in E$ represents user f is user u's follower, and information flow from u to f. From the sub-graph gets node N=702, edge E=1391, average degree=3.963.

Afterwards, we select 10 questions of node degree (0-5) from network, and other 10 questions form node degree (6-10), and 10 more questions from node degree (>11), track their dissemination path. After the questions were released we collect all the information dissemination time and the user ID within one month (30 days).

B. Verify results and analysis

Data obtained by the above-mentioned method can stats daily dissemination nodes, the uninfected node quantity, dissemination nodes calculation, proportion of uninfected nodes and immune nodes of the total nodes. Figure.1 shows trends for the dissemination node *i(t)* changing with time of the user of node's degree between 0-5, 6-10, and >11. Figure.2 shows uninfected nodes *s(t)'s* changing trend with time.

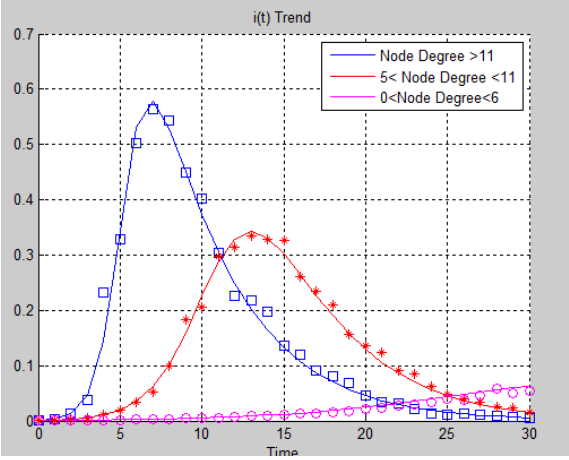

Figure.1 Infected node trend

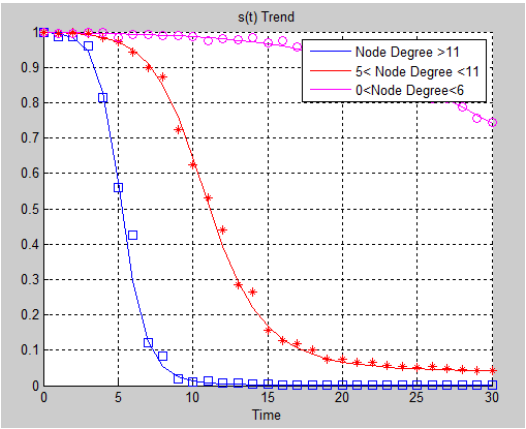

Figure.2 Not infected node trend

As can be seen from Figure.1 and Figure.2, due to high connectivity of ASN, and the threshold of the information dissemination in the network is quite low, even if the number of relationships of initial node is quite little, information can also be spread. When degree of initial node is quite large, information in the network disseminate faster. On the contrary, when degree of initial node is small, information dissemination will show hysteretic nature, i.e. information from no disseminate to start disseminate requires a certain time.

For more precise analysis of the regularity of the information dissemination in the network, we put value *i(t)* and *s(t)* into the model equation, and compute network G's infection probability λ and immune probability μ as follows:

TABLE I: INFECTION PROBABILITY OF λ AND IMMUNE PROBABILITY μ VALUE

Nodes degree	Infection probability λ	Immune probability μ
Node degree between 0-5	0.126	0.328
Node degree between 6-10	0.218	0.257
Node degree >11	0.382	0.216
Average value	0.242	0.267

As can be seen, the higher the node's degree is, the higher probability the information dissemination is; on the contrary, the immune probability doesn't have this trend, and it is consistent with our supposition.

IV. CONCLUSIONS

This paper conducts research into an ASN information dissemination model. Differential equations were established through defining three dissemination regularities and combining with complex network and infectious disease dynamics theory. After obtaining an interpersonal relationship network data sets of an ASN: ResearchGate by programming and analyzing the dissemination behavior of network nodes of different node's degree, results show that due to the high connectivity of ASN, the threshold of the information disseminated in the network is quite low; the higher the degree of initial dissemination node is, the more likely to disseminate information faster in the network is, and the central node has strong influence on society. By the preliminary analysis, ASN can strengthen researchers' relationship and it has a positive correlation.

ACKNOWLEDGEMENTS

This research is supported by Zhejiang Provincial Education Department project (No.: Y201018887).

REFERENCES

[1] Tang, J., et al., ArnetMiner: Extraction and Mining of Academic Social Networks. 2008.

[2] ResearchGate, wikipedia , http://zh.wikipedia.org/wiki/Researchgate, 2012.

[3] Zhang Yan-Chao etc., The research of information dissemination model on online social network. Acta Phys. Sin. Vol.60,No.5(2011): Page66-72.

[4] Newman, M.E.J., The structure and function of complex networks. 2003.

[5] LIU Chang-yu,etc.,Public Opinion Propagation Model Based on Small World Networks[J],Journal of System Simulation,2006-12.

[6] Automata network SIR models for the spread of infectious diseases in populations of moving individuals,N Boccara and K Cheong 1992 J. Phys. A: Math. Gen. 25 2447.

ZHENG Lei , LI Sheng-hong,A Novel Information Diffusion Model based on Microblog Network,Communications Technology,Vol.45 , No.02,2012,Page.39-41