A STUDY OF A INFORMATION PROPAGATION MODEL RELATED WITH TOPIC AND PERSON CHARACTERISTICS

Takao Chiba[a], **Ikuo Suzuki**[a], **Masahito Yamamoto**[a] and **Masashi Furukawa**[a]

[a]Graduate School of Information Science and Technology, Hokkaido University
Kita 14, Nishi 9, Kita-ku, Sapporo, Hokkaido, 060-0814, Japan
{takao, ikuo, masahito, mack}@complex.eng.hokudai.ac.jp

In many studies on the information propagation, a percolation theory and an epidemic model have been used to explain the information propagation. These models make an important role in analyzing how the information propagation affects network strctures and dynamics of the information propagation. However, an individual person is defined uniformly in this model, nevertheless the person interest, one of the attribution of human, has an great effect on the information propagation. This study represents a newly developed information propagation model on the complex network. Then, we perform some experiments to simulate a spread of the topic with our model and compare the propagation pattern of the proposed model with one of the conventional propagation model. In addition, we analyze an easy blogging service on World Wide Web (WWW) to verify effectiveness of our model from a viewpoint of the information propagation.

The information propagation model proposed on the complex network is based on interest of each person. In the complex network, a node and an edge represent a person and the relation between persons, respectively. Our model defines that the topic has a category vector and each person has an interest vector. The topic must at least belong to one category from a set of category $C = \{c_k : (k = 1, \cdots, n)\}$. Thus, topic category vector $\boldsymbol{t}^{(i)}$ is defined as the vector

$$\boldsymbol{t}^{(m)} = [t^{(m)}_{c_1} t^{(m)}_{c_2} \cdots t^{(m)}_{c_k}], \quad \text{where } t^{(m)}_{c_k} = \begin{cases} 1: & \text{topic } m \text{ belongs to the category } c_k \\ 0: & \text{topic } m \text{ doesn't belong to the category } c_k \end{cases}$$

Since each person has defferent interests on topics, we introduce an interest vector $\boldsymbol{a}^{(i)}$ to represent what the person (node v_i) is interested in as follow

$$\boldsymbol{a}^{(i)} = [a^{(i)}_{c_1} a^{(i)}_{c_2} \cdots a^{(i)}_{c_k}], \quad \text{where } a^{(i)}_{c_k} = \begin{cases} 1: & \text{node } v_i \text{ is interested in the category } c_k \\ 0: & \text{node } v_i \text{ isn't interested in the category } c_k \end{cases}$$

According to which category the topic belongs to, the motivation or attention level of the node is determined to spread the information. The category weight vector $\boldsymbol{W}$ represnts the motivation level of category on the information propagation.

$$\boldsymbol{W} = [w_{c_1} w_{c_2} \cdots w_{c_k}] \quad (0 \leq w_{c_k} \leq 1)$$

From the above definition, the propagation power $p_{v_i v_j}$ between node v_i and v_j for topic t_m is introduced as follows,

$$p_{v_i v_j} = \sum_k a^{(i)}_{c_k} a^{(j)}_{c_k} t^{(m)}_{c_k}$$

If there is no shared interest for the category between v_i and v_j, the propagation does not occur in this model.

The propagation proccess is based on the Susceptible-Infected-Recovered (SIR) model on the relation network. However, the infection probability is not uniform. Instead of the infection probability of the SIR model, we use the propagation power $p_{v_i v_j}$ described above. In the propagation proccess, we make a propagation network which consists of nodes recieving the information and edges using as a path of the propagation. Different topics make different propagation networks even if the relation network is same. And if the same topic is propagated, proagation networks vary on the interests of nodes.

We apply our model to analyze a real date "Tumblr (http://www.tumblr.com/)" on WWW, an easy blogging service. In this service, an user can post a content to a user's page and other users can re-post the same content to his own page if he favors it. In other words, an information is propagated by their interests of users. We perform experimental simulations to verify the effectiveness of our model. Moreover we discuss a possibility for predicting the propagation size or paths.

As a consequence, the proposed model has possibilities for the application of simulating or predicting the information propagation, and it is usable for analyzing the information propagation phenomena.

Keywords
blog analysis, boom dynamics, information propagation

References

[1] S.R.Broadbent, J.M.Hammersley, "Percolation process," *Proc.Cambridge Philos.Soc.*, Vol.53, pp.629-641, 1957.
[2] Herbert W Hethcote, "The Mathematics of Infectious Diseases," *SIAM Review*, Vol. 42, No. 4, pp. 599-653, 2000.
[3] Y.Moreno, R.Pastor-Satorras, A.Vespignani, "Epidemic outbreaks in complex heterogeneous networks," *The European Physical Journal B - Condensed Matter and Complex Systems*, Vol. 26, No. 4, pp. 521-529, 2002.
[4] Yamir Moreno, Maziar Nekovee, Amalio F Pacheco, "Dynamics of rumor spreading in complex networks," *Physical Review E*, Vol. 69, No. 6, 066130, 2004.