A New Correlation-based Information Diffusion Prediction

Jong-Ryul Lee
Dept. of Computer Science, KAIST
291 Daehak-ro, Yuseong-gu,
Daejeon, Korea
jrlee@islab.kaist.ac.kr

Chin-Wan Chung
Div. of Web Science and Technology
& Dept. of Computer Science, KAIST
291 Daehak-ro, Yuseong-gu,
Daejeon, Korea
chungcw@kaist.edu

ABSTRACT

For predicting the diffusion process of information, we introduce and analyze a new correlation between the information adoptions of users sharing a friend in online social networks. Based on the correlation, we propose a probabilistic model to estimate the probability of a user's adoption using the naive Bayes classifier. Next, we build a recommendation method using the probabilistic model. Finally, we demonstrate the effectiveness of the proposed method with the data from Flickr and Movielens which are well-known web services. For all cases in the experiments, the proposed method is more accurate than comparison methods.

Categories and Subject Descriptors

H.2.8 [**Database Applications**]: Data mining

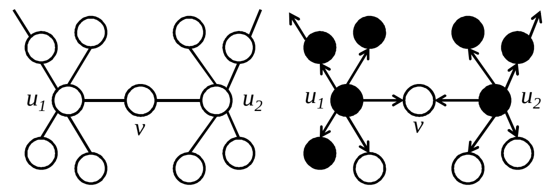
Keywords

Information Diffusion; Recommendation; Social Networks

1. INTRODUCTION

Recently, the number of users using social network services is growing rapidly. For example, Facebook which is one of the most famous social network services has more than 600 million monthly active users. Twitter which is a typical microblogging service has more than 100 million monthly active users. Since there is rich information based on lots of documents and a big network structure in social network services, extracting useful features from large social data became an important issue.

One of the important characteristics that can be extracted from social data is social influence. Social influence occurs when an individual is activated by an action of other people[5, 9]. The activation (a.k.a. adoption) can be any social action shown to other people such as writing and sending a document to friends. Social influence is also a key to explain information diffusion in social networks. Since social influence between two users is based on the social action, it accompanies transferring information from one user to the other. It leads to the fact that if we can predict the occurrence of social influence, we also can predict information diffusion.



(a) Nodes within two hops from v

(b) A snapshot of diffusion process

Figure 1: A part of a social network and a snapshot of a diffusion process on the network

There are several applications of predicting information diffusion. One is the item recommendation. Basically, a recommendation system finds users who are likely to adopt an item and gives those users the item. If an adoption of an item is considered as the activation of a user to the item, it is the same as predicting the next activated user in information diffusion[18]. That is, we can consider that an item is diffused into social networks and the prediction of information diffusion finds users who are likely to adopt the item. In this way, we will apply the prediction to the general recommendation problem in our research.

In this work, to predict the occurrence of social influence, we introduce a new correlation between the activations of users who have an activated common friend in online social networks. When a social network is interpreted as an undirected graph consisting of users(nodes) and relationship links(edges), the new correlation can be found on nodes which share an already activated common neighbor in a social network. For example, Figure 1(a) is a part of a social network and the nodes are within two hops from a node v. Each node represents a user and each edge represents an explicit relationship such as friendship. Figure 1(b) illustrates an intermediate snapshot of a diffusion process in the social network of Figure 1(a). In Figure 1(b), the nodes are the same as those of Figure 1(a), and the edges represent the direction of the diffusion process. More specifically, in Figure 1(b), u_1 and u_2 were initially activated to a document, and five nodes are activated by them. We color those activated nodes as black nodes, and the other nodes, which are not yet activated, as white nodes. In this example, there are four nodes which share u_1 as an activated common neigh-

WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2579241.

bor with v. Suppose that we are trying to estimate the probability that u_1 activates v (i.e., the occurrence of social influence). Since three of them are already activated by u_1, if v and each of the three nodes are positively correlated in terms of being activated by u_1, we can expect that v tends to be activated by u_1. In addition, since the other white node, which share u_1 as an activated common neighbor with v, is not activated, if v and the other node are negatively correlated in terms of being activated by u_1, we can also expect that v tends to be activated by u_1. Depending on the degree of being correlated, we can estimate that the probability that u_1 activates v would be high or low. In this way, for any edge (u, v) in a social network, when u is already activated, to predict the occurrence of social influence from u to v, we can use the correlation between the activation of v and the activations of nodes which share u as an activated common neighbor with v.

Based on this new correlation, we propose a probabilistic model to estimate the probability that a node activates another using the naive Bayes classifier. In addition, we propose a recommendation method using this model, and demonstrate its effectiveness with real datasets.

In this work, we make the following contributions:

- To the best of our knowledge, this paper is the first work to introduce and analyze the correlation between the activations of nodes which share an activated common neighbor.
- Based on the correlation, we propose a probabilistic model to estimate the probability that a node activates another using the naive Bayes classifier. We also propose a recommendation method based on the model.
- We show that the proposed method is more accurate than comparison methods in our experiments.

The rest of the paper is organized as follows. In Section 2, we review related works. We formulate the problem for predicting a user's activation in Section 3. In Section 4, we analyze the new correlation which we found and propose a recommendation method by exploiting the correlation. We demonstrate the effectiveness of the proposed method with real datasets in Section 5. We make conclusions and discuss the future work in Section 6.

2. RELATED WORKS

A lot of research has been conducted on social influence analysis and information diffusion. Anagnostopoulos et al. [1] measure some correlations in social networks based on homophily and confounding. They connect the correlations to their social influence model with a statistical test to show that homophily is related to social influence. Goyal et al. in [7] propose continuous and discrete time models based on exponential decaying to predict the occurrence of social influence.

Some researchers handle mining of topic-level social influence [12, 22, 6, 21, 15, 4, 17]. Dietz et al. [6] propose the citation influence model to calculate the strength of influence between research papers. In addition, a joint latent semantic model is proposed for text and citations in [15]. Liu et al. [12] propose a generative model to mine social influence on a heterogeneous social network. Similarly, Weng et al. [22] use Latent Dirichelt Allocation and hypothesis testing to mine topic-level influence. Chua et al. [4] also propose generative models for item adoptions through exploiting social correlation between users. In contrast to this work, the social correlation is defined between friends on social networks, while we focus on the correlation who share an activated friend. Sang et al. [17] propose a way of mining a topic-sensitive influencer for collaborative recommendations using users' textual annotation and video images.

There are many works for predicting information diffusion[19, 18, 2, 23, 13, 14, 8]. Song et al. in [19] propose a recommendation algorithm based on user's influence to other users with an early adoption based information flow network. However, they assume that the network is homogeneous in terms of the diffusion rate between users. In [18], Song et al. propose an information flow model which leverages interpersonal diffusion rate based on Continuous-Time Markov Chain and apply their model for item recommendation. We compare the method in [18] with the proposed method for demonstration, because it can be directly applied to our problem. Li et al. [11] introduce the concept of the intelligent agent, which jointly considers its interacting neighbors and calculates the payoffs for its different social actions. They propose an information diffusion model using the intelligent agent. Their model determines whether a user will be activated, but does not tell us how likely the user is to be activated. Thus, we do not compare it with the proposed method.

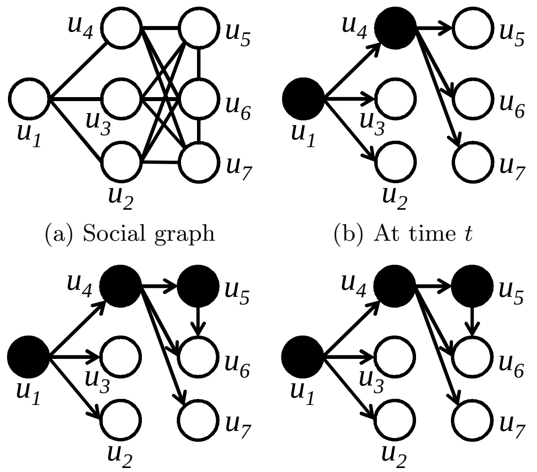
There is a line of research for modeling information diffusion without explicit social links[2, 23, 13, 8]. In [23], Yang et al. introduce a linear influence model to estimate the global influence of a node on the diffusion rate in an implicit network. Yeung et al. introduce implicit user influence from recently activated users to a candidate user regardless of friendship. Matsubara et al. [13] propose an analytical model to predict the rise and fall patterns of information diffusion over time. In addition, Iwata et al. [8] focus on latent influence from sequences of item adoption events and propose a probabilistic model for discovering it.

For information diffusion, empirical studies are also extensively conducted [3, 10, 20]. Sun et al. [20] use Facebook Pages and their associated fans to analyze the mechanics of Facebook Page diffusion. Cha et al. [3] collect the real data from Flickr and analyze it for various features in terms of information diffusion. In this work, we use the dataset from [3] because the existence of social influence in the dataset is proved in [3]. Similarly, Kwak et al. [10] analyze lots of tweets and find interesting features in Twitter such as a non-power-law follower distribution.

From this extensive survey for social influence analysis and information diffusion, there is no existing work for the correlation between users who share an activated friend. We will show the existence and the degree of the correlation, and apply it for item recommendation.

3. PROBLEM DEFINITION

We represent a social graph as an undirected graph $G = (V, E)$ where V is the set of nodes which represent users and E is the set of undirected edges and these edges are mapped to social links between users. For simplicity, we use an undirected graph for modeling a social graph, but our method also works for a directed graph. We denote the set of documents D. A document can be any item shared among



(a) Social graph (b) At time t

(c) At time $t+1$ (d) At time $t+k(k>1)$

Figure 2: A social network and the diffusion steps for document d

users in online social network services such as a photo and an article.

Activation. When a node does social action associated to a document, we say that the node is activated to the document. A node which was already activated to document d cannot be re-activated or inactivated to d. In other words, once a node is activated to document d, then the activation of the node to d is permanent. Instead, the node can be activated to another document $d' \in D$ such that $d' \neq d$. There is an example for the diffusion process of document d over time in Figure 2. Initially, node u_1 introduces document d into the social network in Figure 2(a) before time t. u_4 is activated to d at time t and u_5 is also activated to d at time $t+1$. Lastly, there is no change in Figure 2(d) from the previous status. It means that nobody is activated after time $t+1$.

Activation History. To manipulate the information of users' activations, we define the history of users' activations as a set of tuples (u, d, t) meaning that user u is activated to document d at time t. We denote it as $\mathcal{H}$.

Problem Definition. Given graph $G = (V, E)$, document d, current time t, history $\mathcal{H}$, the problem is to estimate the probability that a node is activated to d.

Using the estimated probability, we will construct an algorithm for item recommendation and demonstrate its effectiveness with real datasets.

4. CORRELATION ANALYSIS AND ITEM RECOMMENDATION

4.1 Correlation Analysis

As we mentioned in Section 2, there are many works which utilize social correlation between the activations of any two nodes or neighbors, but there is no work focusing on the correlation between the activations of nodes who share an activated common neighbor. Thus, let us identify the existence and the degree of the correlation in real data by correlation analysis. For this correlation analysis, we use the Flickr dataset introduced in [3]. In Flickr, there is a function named favorite-marking to express an interest in an item and share it with friends. When a user executes the function to a photo, the photo is marked as favorite by the user and the friends of the user can see the photo. In addition, we use the Pearson product-moment correlation coefficient as a measure of the correlation between the activations of nodes who share an activated common neighbor. For any two nodes $u, v \in V$ and a document d, let $X_{d,v|u}$ denote an indicator variable for the activation of v to d given the activation of u to d. When v is activated to d given the activation of u to d, $X_{d,v|u} = 1$. Otherwise, $X_{d,v|u} = 0$. Given any three nodes $u, v, w \in V$ such that u is a common neighbor of v and w, correlation r between the activations of v and w in terms of being activated by u is computed as follows.

$$r = \frac{\sum_{d \in D_u}(X_{d,v|u} - \overline{X}_{v|u})(X_{d,w|u} - \overline{X}_{w|u})}{\sqrt{\sum_{d \in D_u}(X_{d,v|u} - \overline{X}_{v|u})^2}\sqrt{\sum_{d \in D_u}(X_{d,w|u} - \overline{X}_{w|u})^2}}, \tag{1}$$

where D_u is the set of documents to which u is activated and $\overline{X}_{v|u}$ is the sample mean of variable $X_{d,v|u}$ over $d \in D_u$.

Since there are lots of users in the Flickr dataset, we randomly select 100 users who are activated more than 800 times over the dataset, and denote the set of these users as U. For each selected node u, we randomly pick 100 pairs of u's neighbors (v, w) which are activated more than 100 times over the dataset. Then, we compute r between v and w who share u as a common neighbor. In our analysis, if $r \geq 0.1$, then we say that v and w are correlated. Table 1 illustrates the summary of our correlation analysis. In Table 1, let $Max(r)$ denote the maximum observed value of r. Table 1 says that the maximum value of r is 0.452 among samples, and 89% of all users in U have at least one pair of correlated neighbors in terms of being activated by a common neighbor. In addition, for each node in U, the average number of such pairs is 13.52, and the sample standard deviation is 14.06. Recall that for each user in U, the number of sample pairs of neighbors is 100. Thus, we can expect that about 14% of any two neighbors of a node are correlated with respect to being activated by the node. This result in Table 1

Table 1: Summary of correlation analysis

$Max(r)(\|r\| \geq 0.1)$	0.452
Ratio of users having correlated neighbors	89%
Avg. of # correlated relationships	13.52
S.D. of # correlated relationships	14.06

sufficiently supports the existence of the correlation between the activations of nodes which share an activated common node. Over all samples which have r such that $|r| \geq 0.1$, p-value is lower than 0.0001, when the null hypothesis is that two nodes who share an activated common neighbor are not correlated. It leads to the fact that our correlation analysis is in the high level of confidence.

4.2 Naive Bayes Classifier

4.2.1 A Probabilistic Model

To predict the next activated user, we derive the probability $\pi(u, v)$ that node v is activated by node u based on the correlation between the activations of nodes which share an activated common node. Let us consider that document d is being diffused in graph $G = (V, E)$, time t is the current time, and history $\mathcal{H}$ is given. $\mathcal{H}$ does not store the information of users' past activations, but also the information of users' activations to d before the current time. $\mathcal{H}$ does not store any information of users' activations after the current time. For any two nodes $u, v \in V$ such that $(u, v) \in E$ and u is already activated to d, binary random variable $A_{u\to v}$ is defined as,

$$A_{u\to v} = \begin{cases} 1 & \text{if } v \text{ is activated to } d \text{ by } u \\ 0 & \text{otherwise} \end{cases}. \quad (2)$$

Let $S_{u,v}$ denote the set of nodes who share node u as a common neighbor with node v. We can enumerate the elements in $S_{u,v}$ as list $L_{u,v}$. Let $O_{i|u}$ be a feature variable representing the known state of the i-th node in $L_{u,v}$. If it is observed that the i-th node in $L_{u,v}$ is activated to d before time t after u was activated to d, then $O_{i|u} = 1$. Otherwise, $O_{i|u} = 0$. Given document d, time t, $L = |S_{u,v}|$, and history $\mathcal{H}$, we compute the probability $\pi(u, v)$ using the naive Bayes classification rule as,

$$\pi(u,v) = p(A_{u\to v} = 1 | O_{1|u}, ..., O_{L|u}) \quad (3)$$

$$= \frac{p(A_{u\to v} = 1, O_{1|u}, ..., O_{L|u})}{p(O_{1|u}, ..., O_{L|u})} \quad (4)$$

$$= \frac{p(A_{u\to v} = 1) p(O_{1|u}, ..., O_{L|u} | A_{u\to v} = 1)}{p(O_{1|u}, ..., O_{L|u})} \quad (5)$$

$$= \frac{p(A_{u\to v} = 1) \prod_{i=1}^{L} p(O_{i|u} | A_{u\to v} = 1)}{\sum_{j=0}^{1} p(A_{u\to v} = j) \prod_{i=1}^{L} p(O_{i|u} | A_{u\to v} = j)}. \quad (6)$$

By the definition of the conditional probability, Eq. 4 is derived from Eq. 3 and Eq. 5 is derived from Eq. 4. Then, for simplicity, we assume that given $A_{u\to v}$, each variable $O_{i|u}$ is conditionally independent of another feature variable $O_{k|u}$ for $k \neq i$. This assumption enables us to exploit the observed states of nodes, which share an activated common neighbor with v, without information about how they are activated. Thus, Eq. 6 is derived from Eq. 5 by multiplying each term $p(O_{i|u} | A_{u\to v})$ for $1 \leq i \leq L$.

4.2.2 Parameter Estimation

To use the proposed model derived in the previous section, we estimate parameters for $\pi(u, v)$ with history $\mathcal{H}$ as follows. We denote an estimate of function $f(x)$ where x is a variable as $\hat{f}(x)$. First, for $p(A_{u\to v} = 1)$, we assume that $A_{u\to v}$ follows Bernoulli distribution because the number of the possible outcomes of A_v is 2. Thus,

$$\hat{p}(A_{u\to v} = 1) = \frac{n(A_{v|u} = 1)}{n(A_u = 1)},$$

where $n(A_u = 1)$ is the number of the events that v is activated and $n(A_{v|u})$ is the number of the events that v is activated given the activation of u. In the same way, $p(A_{u\to v} | O_{i|u})$ is computed as follows.

$$\hat{p}(O_{i|u} = 1 | A_{u\to v} = 1) = \frac{n(A_{v|u} = 1, O_{i|u} = 1)}{n(A_{v|u} = 1)},$$

$$\hat{p}(O_{i|u} = 1 | A_{u\to v} = 0) = \frac{n(A_{v|u} = 0, O_{i|u} = 1)}{n(A_u = 1) - n(A_{v|u} = 1)},$$

where $n(A_{v|u} = 1, O_{i|u} = 1)$ is the number of the events that the i-th node in $L_{u,v}$ is activated to the same document to which v is activated, given the activation of u to the document. To compute and store the above probabilities for efficient prediction, we may need a $n \times n \times n$ matrix where $n = |V|$ for storing the information of $n(A_{u\to v} = 1, O_{i|u} = 1)$. Since the space cost for the matrix is too expensive, we estimate it as,

$$\hat{n}(A_{v|u} = 1, O_{i|u} = 1) = \hat{p}(A_{u\to v} = 1) n(O_{i|u} = 1). \quad (7)$$

Since v and the i-th node in $L_{u,v}$ share u as a common neighbor, documents, to which the i-th node is activated given the activation of u, must be also seen by v. By applying $\hat{p}(A_{u\to v} = 1)$ to the number of the documents, we can estimate $n(A_{v|u} = 1, O_{i|u} = 1)$ as Eq. 7. $\hat{n}(A_{v|u} = 0, O_{i|u} = 1)$ can be estimated in the same way. From $\hat{p}(O_{i|u} = 1 | A_{u\to v} = 1)$ and $\hat{p}(O_{i|u} = 1 | A_{u\to v} = 0)$, $\hat{p}(O_{i|u} = 0 | A_{u\to v} = 1)$ and $\hat{p}(O_{i|u} = 0 | A_{u\to v} = 0)$ are easily computed.

4.3 A Recommendation Algorithm

Multiple activated neighbors. We have got the probability that a user is activated by one neighbor at a time. Let us consider the case that a node v has multiple neighbors who are activated to a document. In the case, one of multiple neighbors can activate v. To handle this case, let N_v denote the set of the activated neighbors of v. In general, v will be activated if one of N_v activates v. By assuming that each activated neighbor of v independently activates v from the other neighbors, we can get the probability that v is activated by at least one of N_v. We call it the activation probability $\pi(v)$ of node v and compute it as,

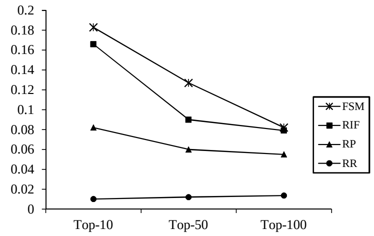
$$\pi(v) = 1 - \prod_{n \in N_v} (1 - \pi(n, v)). \quad (8)$$

Recommendation algorithm. To predict the next activated user of a document, the proposed method calculates the activation probability of candidates, which are not activated yet but have an activated neighbor, and ranks them. The procedure for ranking candidates is illustrated in Algorithm 1. Algorithm 1 ranks candidates to predict the next activated user. In Lines 3-8, we compute the activation probability of each candidate c according to Eq. 8. $activatedNeighbors(c, d, t)$ returns a set of the activated neighbors of c to document d before time t. After the outer loop in Lines 3-8, the algorithm sorts L and returns it.

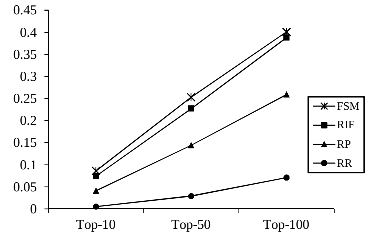
Algorithm 1: rankingProcedure(G, d, t, $\mathcal{H}$, C)

input : G : an input graph, d: an input document, t: the current time, $\mathcal{H}$: the history of users' activations before current time t, C: is a set of candidates
output : L : A ordered list of candidates

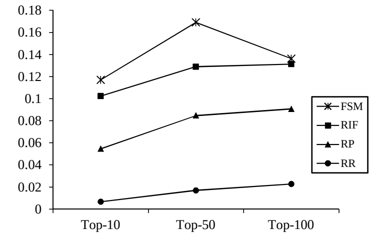
```
1  begin
2      L = [] ;
3      for c ∈ C do
4          π(c) = 1;
5          for n ∈ activatedNeighbors(c, d, t) do
6              π(c) = π(c)(1 − π(n, c);
7          π(c) = 1-π(c);
8          insert (c, π(c)) into L;
9      sort L for the second value of each tuple;
10     return L;
```



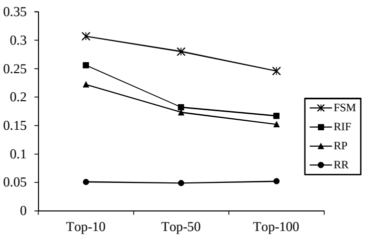
(a) Flickr:Precision



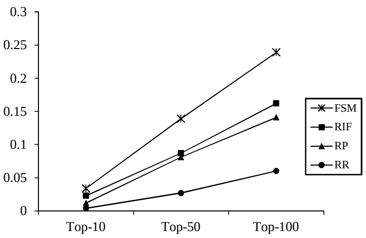
(b) Flickr:Recall



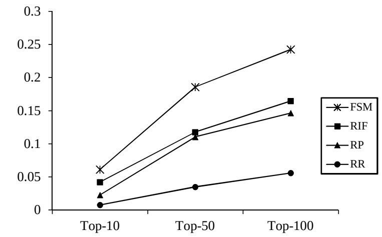
(c) Flickr:F1-score



(d) MovieLens:Precision



(e) MovieLens:Recall



(f) MovieLens:F1-score

Figure 3: The results from the Flickr and MovieLens datasets

5. EXPERIMENT

5.1 Experimental Environment

Comparison methods. We use the method proposed in [18] as a comparison method. The method predicts information diffusion with interpersonal diffusion rate based on Continuous-Time Markov Chain and can be applied directly to our problem. In addition, we use two naive methods which are a random recommendation method and a random probability method for comparison. The random recommendation method retrieves random users as a result. The random probability method is the same as the proposed method, except a random value is assigned to $\pi(v)$ for $v \in V$. In this experiment, we label the proposed method as FSM (Friend Sharing relationship-based Model), the method in [18] as RIF (Rate-based Information Flow model), the random recommendation method as RR and the random probability method as RP.

Datasets. To demonstrate the performance of the proposed recommendation algorithm, we choose Flickr and MovieLens as datasets. The Flickr dataset comes from [3]. There are about two million users and thirty million friendship links in the Flickr dataset. For the experiments, we reduce the dataset from Flickr by selecting the users who are activated to more than 800 photos. After reducing the dataset, the number of users becomes 5,926 and the number of edges becomes 675,124. We still have 13,867,984 favorite-markings and 11,267,320 photos. The MovieLens dataset used in [18] consists of 6,040 users and 1,000,209 ratings[16]. Since there is no explicit relationship in MovieLens, we generate an explicit link between two users when the number of ratings done by the two user to the same items is more than 200. The number of generated edges for MovieLens is 1,020,797. In addition, each user has at least 20 ratings and we assume that a rating corresponds to an activation in this work. We divide the history of each dataset into the training data and the testing data in terms of time. The 90th percentile of each dataset is used as a training set and the other is used as a testing set.

5.2 Experimental Results

For this experiment, we randomly select 100 documents which are adopted more than 30 times in the Flickr dataset and 10 times in the MovieLens. Since FSM, RIF and RP are designed to only predict users who will be activated and have an activated neighbor, we filter out users who do not have any activated neighbor from an answer set.

Top-k Test for Recommendation Performance. In this top-k test, for each randomly selected document in the Flickr dataset, we find the time when the 40th percentile of all activated users were already activated, and then make the comparison methods predict users who will be activated after the time. For the MovieLens, we use the 10th percentile.

Figure 3 illustrates the results of the precision, recall and F1-score tests over the Flickr and MovieLens datasets. In Figure 3, the recommendation performance of the proposed method is better than those of the other methods in most cases. In the Flickr dataset, FSM averagely improves F1-score by 16% compared to RIF. Especially, FSM improves F1-score by 31% compared to RIF and precision by 41% compared to RIF in the top-50 test. In the MovieLens dataset, FSM averagely improves F1-score by 50% compared to RIF. FSM improves precision by 53%, recall by 59%, and F1-score by 58% compared to RIF in the top-50 test. The performance gaps between FSM and each of the two random-based methods are even bigger.

In most cases, FSM has better recommendation performance than the other methods. Thus, we identify the effectiveness of the proposed method and the correlation between the activations of nodes which share an activated common neighbor.

6. CONCLUSIONS

In this paper, we study the new correlation between the activations of users who share an activated friend in online social networks. Based on the study of the correlation, we formulate the naive Bayes classifier by estimating the probability that a node is activated by another given the observed states of nodes which share an activated common neighbor with the node. Finally, we construct the recommendation method using the classifier and perform the experiments to demonstrate the effectiveness of the proposed method.

In the future, we will extend our study for the correlation which we found. For example, we can extend the correlations analysis by considering a correlation between users who participate in the same community. In addition, we will consider another classification algorithm to more effectively exploit the correlation which we studied.

7. ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea grant funded by the Korean government (MSIP) (No. NRF-2009-0081365).

8. REFERENCES

[1] A. Anagnostopoulos, R. Kumar, and M. Mahdian. Influence and correlation in social networks. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, pages 7–15, 2008.

[2] C.-m. Au Yeung and T. Iwata. Capturing implicit user influence in online social sharing. In *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia*, HT '10, pages 245–254, 2010.

[3] 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*, WWW '09, pages 721–730, 2009.

[4] F. Chua, H. Lauw, and E.-P. Lim. Generative models for item adoptions using social correlation. *Knowledge and Data Engineering, IEEE Transactions on*, 25(9):2036–2048, 2013.

[5] R. B. Cialdini. Influence: science and practice. 1993.

[6] L. Dietz, S. Bickel, and T. Scheffer. Unsupervised prediction of citation influences. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pages 233–240, 2007.

[7] A. Goyal, F. Bonchi, and L. V. Lakshmanan. Learning influence probabilities in social networks. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, WSDM '10, pages 241–250, 2010.

[8] T. Iwata, A. Shah, and Z. Ghahramani. Discovering latent influence in online social activities via shared cascade poisson processes. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '13, pages 266–274, 2013.

[9] R. E. Kraut, R. E. Rice, C. Cool, and R. S. Fish. Varieties of social influence: the role of utility and norms in the success of a new communication medium. *Organization Science*, 9:437–453, April 1998.

[10] H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, pages 591–600, 2010.

[11] D. Li, Z. Xu, Y. Luo, S. Li, A. Gupta, K. Sycara, S. Luo, L. Hu, and H. Chen. Modeling information diffusion over social networks for temporal dynamic prediction. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*, CIKM '13, pages 1477–1480, 2013.

[12] L. Liu, J. Tang, J. Han, M. Jiang, and S. Yang. Mining topic-level influence in heterogeneous networks. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, pages 199–208, 2010.

[13] Y. Matsubara, Y. Sakurai, B. A. Prakash, L. Li, and C. Faloutsos. Rise and fall patterns of information diffusion: Model and implications. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, pages 6–14, 2012.

[14] S. A. Myers, C. Zhu, and J. Leskovec. Information diffusion and external influence in networks. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, pages 33–41, 2012.

[15] R. M. Nallapati, A. Ahmed, E. P. Xing, and W. W. Cohen. Joint latent topic models for text and citations. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, pages 542–550, 2008.

[16] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. Grouplens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, CSCW '94, pages 175–186, 1994.

[17] J. Sang and C. Xu. Social influence analysis and application on multimedia sharing websites. *ACM Trans. Multimedia Comput. Commun. Appl.*, 9(1s), Oct. 2013.

[18] X. Song, Y. Chi, K. Hino, and B. L. Tseng. Information flow modeling based on diffusion rate for prediction and ranking. In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, pages 191–200, 2007.

[19] X. Song, B. L. Tseng, C.-Y. Lin, and M.-T. Sun. Personalized recommendation driven by information flow. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '06, pages 509–516, 2006.

[20] E. Sun, I. Rosenn, C. Marlow, and T. Lento. Gesundheit! modeling contagion through facebook news feed. *Proc. ICWSM*, 9, 2009.

[21] J. Tang, J. Sun, C. Wang, and Z. Yang. Social influence analysis in large-scale networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '09, pages 807–816, 2009.

[22] J. Weng, E.-P. Lim, J. Jiang, and Q. He. Twitterrank: Finding topic-sensitive influential twitterers. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, WSDM '10, pages 261–270, 2010.

[23] J. Yang and J. Leskovec. Modeling information diffusion in implicit networks. In *Proceedings of the 2010 IEEE International Conference on Data Mining*, ICDM '10, pages 599–608, 2010.