A Diffusion Mechanism for Online Advertising Service Over Social Media

Yung-Ming Li and Ya-Lin Shiu

Institute of Information Management, National ChiaoTung University, Hsinchu, Taiwan

Abstract. Social media has increasingly become a popular platform for diffusing information, through the message sharing of numerous participants in a social network. Recently, companies attempt to utilize social media to expose their advertisements to appropriate customers. The success of message propagation in social media highly depends on the content relevance and closeness of social relationships. In this paper, considering the factors of user preference, network influence, and propagation capability, we propose a social diffusion mechanism to discover the appropriate and influential endorsers from the social network to deliver relevant advertisements broadly. The proposed mechanism is implemented and verified in one of the most famous micro-blogging system-Plurk. Our experimental results shows that the proposed model could efficiently enhance the advertising exposure coverage and effectiveness.

1 Introduction

In recent years, social media has been flourished and raised high much popularity and attention. Social media provides a great platform to diffuse information through the numerous populations. Common social media marketing tools include Twitter, Youtube, Facebook and so on. An overwhelming majority (88%) of marketers are using social media to market their businesses, and a significant 81% of marketers indicate that their efforts in social media have generated effective exposure for their businesses, according to a social media study by Michael Stelzner[12].In 2010, one of the most popular micro-blogging websites, Twitter, announced an innovative advertising model, " Promoted Tweets". Promoted Tweets makes tweets as ads, which are distinctive from both traditional search ads and recent social ads. They measure the advertising performance and the payment of sponsored tweets by "resonance" - the interactions between users and a particular sponsored tweet such as retweet, reply, or bookmarking [13]. How to choose the right people to deliver the information, how to take the advantage of the social media, and how to design an ads diffusion mechanism to widen the spreading coverage are crucial issues in the online advertising campaign. In this paper, to address these issues, we design a social media diffusion mechanism, based on the concepts of content recommendation and network routing. Once the appropriate messages and diffusion paths are identified, message can be effectively delivered with support of the generated information. Considering the factors of user preference, network influence, and propagation capability, our system can

Li Y. and Shiu Y. (2010).
A Diffusion Mechanism for Online Advertising Service Over Social Media.
In *Proceedings of the 4th International Workshop on Architectures, Concepts and Technologies for Service Oriented Computing*, pages 42-51
DOI: 10.5220/0003049500420051

effectively identify the most appropriate nodes in the social network for delivering the relevant ads and recommend the friends for information sharing for an intermediate node.

2 Related Literature

2.1 Social Media

Social media are Internet platforms designed to disseminate information or messages through social interaction, using highly accessible and scalable publishing techniques. Social media is composed of content (information) and social interaction interface (intimate community engagement and social viral activity). With its emerging trend and promising popularity, researchers have put academic efforts in analyzing the characteristics and functionalities of social media. For example, Kaplan and Haenlein [6] examine the challenges and opportunities of social media and recommend a set of ten rules that companies should follow when developing their own social media strategy. To effectively communicate with customers, researchers engaged in analyzing marketing trends and social relations. Gilbert and Karahalios [4] develop a predictive model that maps data of social activity to tie strength so as to improve design elements of social media.To better figure out the users' behaviors, many researchers analyze the social influence, social interactions, and information diffusion in social media[3].Comparing to the existing works, the study of information diffusion mechanism design of social media is apparently rare and new.

2.2 Online Advertising

The issue of online advertising has aroused much academic interests and been spotlighted for decades. Online advertising usually could be categorized into two types: 1) targeting advertising, which deliver the ads based on user's preference profiles, 2) social advertising, which deliver the deliver the ads the influential users determined by social relationship[11]. Targeted advertising usually applies the content-based and collaborative-based approaches to discover users' personal preferences. Compared with the traditional online advertising, social advertising is a form of advertisement that addresses people as part of a social network and uses the social relations and social influences between people for selling products [14]. Some researchers measure the influential strength by analyzing the number of network links and users' relation and interaction in the network to identify the influential nodes for social advertising[11, 14]. In this paper, considering the factors of user preference, network influence, and propagation capability, we propose a social diffusion mechanism to identify the appropriate and influential endorsers from the social network to deliver relevant advertisements broadly.

2.3 Information Diffusion and Social Routing

Researchers analyze information diffusion in the social network based on individual's characteristics. Some based on the bond percolation, graph theory or probabilistic model to extract the influential nodes, considering the aspect of dynamic characteristics, such as distance, time, and interaction and so on [7-8]. By revealing influential factors and realizing the processes of the information diffusion, marketers can predict when and how the information spreads over social networks to maximize the expected spreading result [5]. In this paper, we include static and dynamic factors dimensions to evaluate the propagation ability of nodes in social network.The design of social diffusion mechanism is conceptually similar to that of computer network routing process in selecting paths to switching the packet. Routing directs packets to be forwarded from their source toward their ultimate destination through intermediate nodes; hardware devices usually called routers. However, in the context of social network, the links in social networks are formed by social relations and interaction and researchers focus on the study of the issue: delivering the right information to the right nodes and spreading widely. The goal can be achieved by implementing feasible approaches to discover the influential nodes and leveraging the social relations to diffuse the information between users further.In our paper, we incorporate the concept of network routing to develop a social endorser engine to generate "social routing tables" to support the information diffusion in a social network.

3 System Architecture

Analogous to the routing process in computer networks, we design a social diffusion mechanism which sends a recommended list of users to our initial nodes with suitable path for information diffusion. The recommendation lists suggest the users who have strong propagation abilities in social networks. The users are referred as social endorsers potentially willing to transmit the information to all his/her friends.

Notice that the proposed social diffusion mechanism is different from spamming. We recommended those friends based on their preferences, social influence, and propagation abilities via quantitative measurement. The advertising message will be guided to right people by user's judgments with the support of the system recommendation. If users deliver ads to their friends, it means that users also think their friends like the ads. The mechanism takes the advantage of content relevance and social relation to reduce the negative impression of the advertisement and gain the advertising effectiveness. Social media provided us the source data of individual's preference, social relation and social influence. Preference is an important issue in target advertising. The social influence between users of the social media happened when they are affected by others. It is likely that we are usually influenced by our friends or our family. The social relation is a crucial factor to empower the social influence. If a user frequently interacts with someone, to some extent, there are more similarities and closeness between them. Therefore, we incorporate these components into our proposed social endorser discovery engine.

3.1 Social Endorser Discovery Engine

Effective information diffusion on social networks is grounded on the relevance of individual preference and closeness of social relations. Therefore, the main functionality of the proposed social endorser discovery engine is to identify the nodes with the strong propagation ability in disseminating relevant messages as wider as possible. In order to identify the appropriate social endorsers so as to achieve a better diffusion performance, in this research, we not only consider the static factor, but also dynamic factor in the evaluation of nodes' diffusion capability - transmit information towards the most suitable friends and spread widely further. Figure 1 displays the main components and procedures of our social endorser discovery engine.

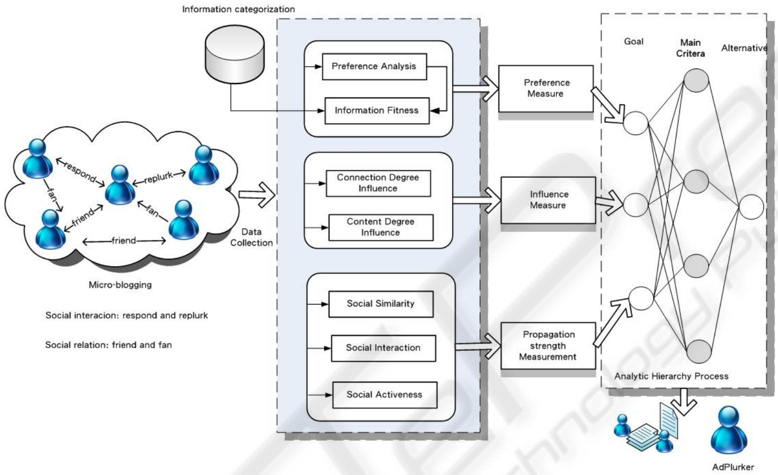

Fig. 1. Architecture of endorser discovery engine.

User Preference Analysis Module

As people trend to share information interesting to the receiver, discovering users' preference is an important factor to be considered in the social endorser discovery. By analyzing users' preference, we can better understand what kinds of information are suitable to be shared between the users. In order to realize user's affinity levels of information categories, we adopt a tree-like structure to categorize a set of information and users' preference. Tree-like structure is practically employed in many researches, such as product taxonomy [1, 9].Besides, we utilize a distance-based approach to calculate the similarity between user categories and information categories. The preference of a user and the type of an advertising message are described by a catalog node they belong. Assume and stand for the category 1 (a user's preference) and catalog 2 (an advertisement type) belong to respectively and represent the catalog of the first mutual parent nodes of catalogs 1 and 2. The fitness degree of the ads to a user can be calculated by the following formula:

$$Sim_P\left(C_1,C_2\right)=\frac{2D_{fn}}{D_1+D_2+2D_{fn}} \tag{1}$$

Network Influence Analysis Module

Connection Degree Influence. For the purpose of evaluating the relative importance of user position in the whole network, social network analysis is applied. Degree centrality is defined as the number of direct connections/links upon a node. Specifically, in-degree is a count of the number of connections directed to the node, and out-degree is the number of connections that the node directs to others. In this research, first, we consider the spammer or bots attempt to follow many people in order to gain attention. Secondly, based on in-degree or out-degree ignores the ability for a user to interact with content in the social network. Therefore, we use mutual relation (friendship) to measure the degree centrality as in practice, mutual degree represents the number of friends a user has. Mutual degree for user i is measured as

$$MD(i) = \sum_{j=1}^{n} E_{ij} \tag{2}$$

where is 1 if an edge from node i to j exists and an edge from node j to node i exists, too, otherwise it is 0.

Content Degree Influence. Content Degree Influence is used to evaluate the popularity of what a user posts. We measure the content degree of a user by the count of the total responses and message forwards by people. We denoted as the total number of the elements in a set. The formula for content degree measure can be expressed as:

$$CD(i) = \frac{|\Phi_{reponse(i)}| + |\Phi_{forwrad(i)}|}{|\Phi_{post(i)}|} \tag{3}$$

where $\Phi_{post(i)}$ stands for a set of the messages posted by user i, $\Phi_{reponse(i)}$ represents the set of the responses on user i 's posts , and $\Phi_{forward(i)}$ is the set of i's posts forwarded by others. The aggregate network influence score is the sum of the mutual degree value $MD(i)$ and the content degree value $CD(i)$.

Propagation Strength Analysis Module

Social Similarity. Social similarity aims to measure the similarity of two people from implicit social structure and behaviors, such as “friend-in-common” and “content-in-common”. The more friends-in-common of two people generally reflects the higher connection level between them. If two people have more common friends, their interests should be more similar and the possibility that a people will forward a message he feels interesting to the other becomes higher. Denote $F(i)$ as a set of user i 's friends. The similarity of friend-in-common between user i and friend j , is measured as:

$$Sim_{CF}(i,j) = \frac{|F(i) \cap F(j)|}{Max(|F(i)|,|F(j)|)} \tag{4}$$

In addition, the more content-in-common posted by two people, the higher similarity degree between them. Semantic analysis can be use to evaluate the social similarity measure in the aspect of content-in-comment and to discover potential preference of users [2]. Specifically, traditional information retrieval (IR) technology can be used to analyze the semantics of content. To examine the semantic similarity among posts, we use CKIP Chinese word segmentation system to parse and stem the crawled contents and apply the analysis of frequency–inverse document frequency (TF-IDF) weight to measure how important a word is to a document in a collection or corpus.

$$tf_{i,j} = \frac{freq_{i,j}}{\max_l \left(freq_{l,j} \right)} \tag{5}$$

where $freq_{i,j}$ as the raw frequency of term i appear in post j and $\max_l \left(freq_{l,j} \right)$ is the number of times the most frequent index term, l, appears in post j . The inverse document frequency for term i is formulated as

$$idf_i = \log \frac{N}{n_i} \tag{6}$$

where N is the total number of posts and n_i is the number of posts in which the term i appears. The relative importance of term i to post j can be obtained by calculating

Then, we measure the similarity degree between people by a cosine similarity metric. The similarity of corpus between user *i* and friend *j* is defined as:

$$Sim_{CC}(i,j) = \cos\left(\vec{i}, \vec{j}\right) = \frac{\vec{i} \cdot \vec{j}}{|\vec{i}\,||\,\vec{j}\,|} \tag{7}$$

where $\vec{i}$ and $\vec{j}$ are the two vectors in the *m* dimensional user space which is the keywords to a person in a collection or corpus.

Finally, the total social similarity (SS) score is the sum of “friend-in-common” and “content-in-comment” value.

Social Interaction. Social interaction is different from social similarity since social interaction explicitly catches the factor of dynamic actions between people. It can be used to evaluate the intimacy between two users. For instance, it is common for a user to respond or forward someone’s message. It is reasonable to assume that more interaction activities would lead to a higher probability to transmit the information as they are usually interested in mutual messages. Given two users *i,j* ,social interaction between them can be formulated as:

$$SI(i,j) = \frac{|\Phi_{response(i,j)}|}{|\Phi_{response(i)}|} + \frac{|\Phi_{forward(i,j)}|}{|\Phi_{forward(i)}|} \tag{8}$$

$\Phi_{response(i,j)}$ is the set of responses generated by user *j* to user *i*’s posts and $\Phi_{forward(i,j)}$ is the set of forwards conducted by user *j* to user *i*’s posts.

Socail Activeness. Social activeness is used to calculate the activity intensity of a user. A user with higher activeness indicates a lager level the user is engaged in information sharing or discussion with others and a higher probability to transmit the information. We calculate the activeness of a user by the count of post records during a period of time in a social platform. The formula is defined as below:

$$SA(i) = \frac{\sum_{i=1}^{T} \left| \Phi_{messages(i,t)} \right|}{T} \tag{9}$$

where $\left| \Phi_{messages(i,t)} \right|$ is the total number of messages posted by user i at time period t.

The propagation strength measurement is used to evaluate the user whose network propagation capability of a user. The propagation strength of a user is measured by aggregating the propagation strength and can be computed in a recursive way. To enhance advertising efficiently, it is important to pay attention on next layer's propagation capability. Though advertisers delivered advertisements to a social endorser with high propagation capability, they can't ensure the social endorser's friends with high propagation capabilities equally. Therefore, we thought friends' propagation capabilities would affect a social endorser's propagation capability. Friends' propagation capabilities became a dimension to measure a social endorser propagation capability. In other words, individual's propagation capability is affected by their friends.

The propagation strength is formulated as below:

$$PA(i) = SA(i) + \sum_{j \in F(i)} PA(j) \cdot \left(SS(i,j) + SI(i,j) \right) \tag{10}$$

where $F(i)$ denotes a set of user i's friends.

4 Experiments

Micro-blogging service has become one of the top tools for social media marketing. Compared to traditional blogging, micro-blogging allows users to publish brief messages make people easy to read and repost.These characteristics: brief messages, instant, easy to read and easy to share make micro-blogging become a good platform to conduct social media marketing Therefore, in this research, we apply and validate our proposed mechanism in micro-blogging systems. We conduct our experiments in Plurk.According to Alexa, 2010, the user of Plurk is more than Twitter and 34.4% of Plurk's traffic comes from Taiwan. Users of Plurk can connect with their friends via lots of functions such as updating instant messages, sharing image or video to your friends and responding friends' messages. Besides the well-constructed network structure, another important reason of choosing Plurk as our platform is they provide the APIs for developers to easily request the data of users and networks which is helpful to crawl more complete data to conduct our experimental work.

4.1 Data Description

In our mechanism, AdPlurker will send private messages with ads to users whom we discover by different approaches. Users who receive the messages, which include brief information of the ads and a recommended list of users who are also interested in the ads and have higher propagation capabilities in their network. User can click the hyperlinks of brief information to get detailed information and the click-through record will be collected for evaluation work. Also, users can share the messages with their friends who are recommended by the system. The transmitted message records will be collected. To preventing click fraud, we recorded one click for each individual user for the same advertisement. We conduct the experiment during the period of 11 April to 13 May.

In the preference module, we collect target users' explicit preferences by questionnaires. Besides, in order to better realize users' preferences, we collect implicit preferences from the behaviors in Plurk. In the Plurk, "become fans" is a function for users to follow others' plurks and also declare their preferences for information type. We use these data to match with the hierarchy of product category of Rakuten, one of the famous online shopping mall, to structure the preference category tree of each user. In the influence module, the out-degree measure and social popularity are taken into consider. The friend links usually are the strongest links and imply the structural influential in the network. A user is attention-getting since his/her plurks is often replurked and responsed by others. It also means he/she is influential in content. In propagation strength measurement, we analyze users' occurring activities during the recent six months: daily plurks, responses and replurks as the active and social interaction measure, the similarity in friends and content as the social similarity measure. We calculate the statistics of these as the propagation strength measurement.

We develop a Plurk robot named as AdPlurker and invite users who are active and have used Plurk for a long time to join the experiments. Until April 2009, There are 107 users (55% male, 45% female) aging between 20-50. To simulate a real network structure, our target users are formed with different locations and careers. There are 121,837 users and 971,014 plurks in our database. We collect data from the target users to 3rd-4rd layers, due to the degrees of separation is limited the layer to 3rd-4rd layer[10], and crawl their plurks, responses, and interactions with friends that happened within six months.

4.2 Experimental Results

In order to evaluate the performance of our proposed mechanism, we used the click-through rate (CTR) [20] and repost-through rate (RTR) [13]. The former is a practical statistics about advertising effectiveness; the latter is an effective means to evaluate the eventual spread of the advertisement. Also, the two performance indicators are the key measures in promoted tweets which is newly social advertisement platform proposed by Twitter. The CTR formula is defined as:

$$CTR = \frac{\Phi clicks}{\Phi delivered} \tag{11}$$

where $\Phi clicks$ is is the total number of clicks and $\Phi delivered$ is total number of ads delivered. The RTR formula is defined as:

$$RTR = \frac{\Phi repost}{\Phi delivered} \tag{12}$$

where $\Phi repost$ is the total number of repost and $\Phi delivered$ is total number of ads delivered. We compare four online advertising approaches, which are commonly used in micro-blogging. These different approaches are described as follows.

In-degree. It is the most common measure used to evaluate the influence of micro blogging by the number of fans. This measure is currently employed by many other third-part services, such as twiiterholic.com and wefollow.com

Ratio-degree. The measure is similar to the ratio between the number of a user's followers and the number of other people that the user follows. It was proposed from the Web Ecology Project, an interdisciplinary research group based in Boston, Massachusetts.

Preference + Out-degree. Discovering the topic-influential nodes for delivering advertising message by taking the advantage of the target advertising and social influence.

Social Diffusion. The approach we proposed in this study. We applied analytic hierarchy process (AHP) to realize the final weight combinations of three components.

Figure 2 shows the results of different advertising strategies. According to the database, the advertisements of in-degree approach got total 0.157 CTR in 2042 delivery times. Ratio-degree approach got 0.176 CTR in 3922 deliveries. The hybrid approach of preference and out-degree got 0.217 CTR in 4596 deliveries. Our social diffusion mechanism got 0.299 CTR in 7356 deliveries. Comparing the diffusion performance in the four advertising approaches, we can observe that our proposed social diffusion approach has the best coverage and exposure in advertising campaign.

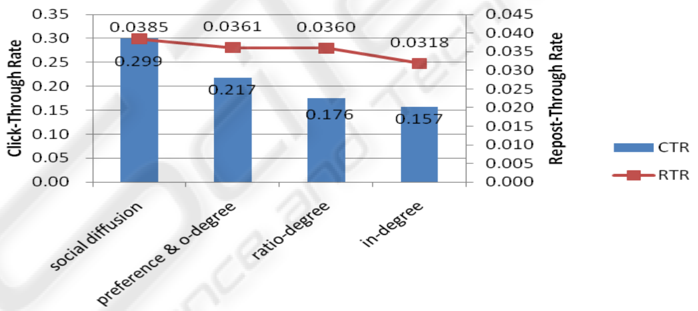

Fig. 2. Performance comparisons of various endorser discovering strategies.

5 Conclusions

In this paper, we propose a social diffusion mechanism to discover the nodes with the strong propagation capability in delivering advertising information and recommend each intermediate node a list of nodes with the high prior propagation so as to enhance the efficiency and effectiveness of spreading advertising message. We combine the static factor, which includes individual preference and link structure of relationship and the dynamic factor, which includes social interactions and social similarity between the nodes, to develop our model. Our experimental results get positive outcomes in both click-through rate and repost rate, and reveal some implicit connections between the components in the framework. A better CTR reflects that our mechanism can raise the visibility of advertising information. And a higher RTR indicates a higher exposure of the advertising and reveals that users are interested in the advertisement shared by friends and willing to share them with others. Our proposed mechanism can widely extend the diffusion coverage of ads. It provides the advertising sponsors a powerful vehicle to successfully conduct advertising diffusion campaigns.

References

1. Albadvi, A., Shahbazi, M. A hybrid recommendation technique based on product category attributes. Expert Syst. Appl.(2009).
2. Berendt, B., Navigli, R. Finding your way through blogspace:Using semantics for cross-domain blog analysis. 2006.
3. Delre, S. A., Jager, W., Bijmolt, T. H. A. and Janssen, M. A. Will it spread or not? The effects of social influences and network topology on innovation diffusion. Journal of Product Innovation Manageme(2010).
4. Gilbert, E., Karahalios, K. Predicting tie strength with social media. In Proceedings of the Proceedings of the 27th international conference on Human factors in computing systems(2009) ACM.
5. Iribarren, J. L., Moro, E. Information diffusion epidemics in social networks. Physics and Society(2007).
6. Kaplan, A. M., Haenlein, M. Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 59-68.
7. Kempe, a., Kleinberg, J.,Tardos, É. Influential Nodes in a Diffusion Model for Social Networks. Springer Verlag2005.
8. Kimura, M., Saito, K., Nakano, R., Motoda, H. Finding Influential Nodes in a Social Network from Information Diffusio. Springer US 2009.
9. Leung, C. W.-k., Chan, S. C.-f.,Chung, F.-l. A collaborative filtering framework based on fuzzy association rules and multiple-level similarity. Knowl. Inf. Syst.(2006) 357-381.
10. Li, Y.-M., Chen, C.-W. A synthetical approach for blog recommendation: Combining trust, social relation, and semantic analysis. Expert Systems with Applications(2009) 6536-6547.
11. Li, Y.-M., Lien, N.-J. An endorser discovering mechanism for social advertising. In Proceedings of the Proceedings of the 11th International Conference on Electronic Commerce(2009) ACM.
12. Stelzner, M. Social Media Marketing Industry Report. 2009.
13. Twitter Promoted Tweets. 2010.
14. Wen, C., Tan, B. C. Y. and Chang, K. T.-T. Advertising Effectiveness on Social Network Sites: An Investigation of Tie Strength, Endorser Expertise and Product Type on Consumer Purchase Intention. 2009.