A Diffusing Path Planning Mechanism for Marketing Information Propagation over Social Media

Yung-Ming Li
Institute of Information Management, National Chiao-Tung University, Taiwan
yml@mail.nctu.edu.tw

Cheng-Yang Lai
Institute of Information Management, National Chiao-Tung University, Taiwan
cylai.k@gmail.com

Abstract

Social media are gaining importance as a component of marketing strategies. Many of them, such as social networking site, blogosphere, and micro-blogosphere, have been seeking business opportunities and establishing brand expression in recent years. Online marketing information diffusion becomes the critical business model for online social networks. Marketers attempt to diffuse advertisements for reaching potential customers through the Internet. However, current marketing research discovered possible influencers but not appropriate support them to diffuse advertisements. In this research, a diffusing path planning mechanism for advertisement was developed for supporting influencers to propagate marketing information and supporting marketers to evaluate possible reward under different marketing strategies.

1. Introduction

A previous study by DEI Worldwide [8] conducted an online survey across a national sample in United States and highlighted that 70% of consumers visit social media for getting information; 49% of these consumers make purchase decisions based on the information they got from social media; 60% of consumers would like to pass along information to friends through social media; and 45% of consumers engaged in word-of-mouth within social media. The report states that companies would lose lots of business opportunities if they are not engaged in social media as a part of their online marketing strategy. Another business report [30] also reported that advertising through online media has increasing year by year greatly (22% increase from 2010 to 2011). Also, according to the report by Stelzner [31] and Nielsen’s [29], there are 83% of marketers who said that social media was important to their businesses and 70% of social media users engage in online shopping. Obviously, the use of social networking media for business marketing purposes is not a suggestion but necessity.

Internet-based marketing, delivering marketing information over the Internet, has become one of significant promotion methods for businesses. Recently, sellers (enterprises and individuals) have promisingly turned to propagate marketing information through online media for seeking business opportunities (e.g. advertisements) [24,34,35] and for establishing brand expression (e.g. branding message) [17,20,21]. In recent years, social media such as social networking sites (e.g. Facebook), blogospheres (e.g. Blogspot), and micro-blogospheres (e.g. Twitter and Plurk) provide powerful means of organizing friend network, publishing contents, and sharing information [26]. They have been massively developed for business usage, such as viral marketing and online advertising.

Information diffusion through online social networks has recently become an active research topic [1]. Generally, to our knowledge, information diffusion related research is applying related analysis (e.g. social network analysis) to identify the influencers or endorsers who might most help to diffuse [7,16,36]. Influencers or endorsers commonly selected through recommender systems, which are expected to reach and influence potential customers [18,25]. However, which path these influencers/endorsers should pass the marketing information to is not well known. That is, which path is the best diffusion path if the information to be diffused starts from him/her while an evaluated influencer receives marketing information?

Since the Internet is become one of the primary message delivering medium, it is worthwhile to investigate and design a novel mechanism for supporting online marketing information propagation. Specifically, the aim of our research is to analyze the optimal diffusion path that could:

1. assist the marketers to evaluate possible diffusion effectiveness under different...
marketing strategies (e.g. raising product sales or brand awareness);
(2) assist the evaluated influencers or endorsers to propagate information to specific individuals for gaining required diffusion reward.

This paper is organized as follows. The brief descriptions of the basic concepts related to our research topics are provided in Section 2. In Sections 3, the model combined with SNA, Markov chain and optimization computation is proposed. A practical experiment is given in Section 4 and the experiment is given in Section 5. In Section 6, some concluding remarks are provided.

2. Literature Review

2.1. Social Media Marketing

Social media marketing is a new and rapidly growing way in which businesses are reaching out to the potential customers. Jackson [15] has shown that online media are more effective in influencing consumers than direct marketing channels. The use of social networks allows companies to engage with customers to a degree that outpaces traditional advertising.

Social media marketing embraces many possible techniques for advertising and branding across social networks such as social networking site, blogosphere, and micro-blogosphere [32]. Social media marketing has become such one of important features that it is no longer a question of whether to use it, but how to use it [19]. Therefore, in this study, we proposed a mechanism for planning the diffusion path for dealing with the advertising and branding information diffusion problem.

2.2. Information Propagation

Information propagation in online social sites has attracted great research interests recently. Information diffusion techniques in social networks are broadly used for influencing and informing people [11]. The positive effects of viral marketing to influence [22], and word-of-mouth [12] to inform potential consumers have been observed. Obviously, an information (e.g. branding information) informed by friends is more trustworthy and acceptable than marketers [23].

Peer influence means that an individual might lead other individuals to do things according to the information gathered from him/her. Prior work has shown that peer influence has a positive effect in online marketing [6,9,33]. In [36], the authors developed a linear influence model to predict the possible influential nodes in the network for modeling the information diffusion in online social media. Kempe et al. [16] proposed an algorithm which considered the maximum social influences in social network to find the minimum set of influencers.

However, these models do not take the problem of diffusion path planning into account. In this research, the information diffusion problem is seen as a sequential path planning optimization problem rather than a recommender problem. So we proposed a simple maximization model for planning the optimal diffusion path for the influencers who are evaluated by other mechanisms.

2.3. Social Network Analysis

A social network, described by a graph of relationships and interactions within a group of individuals, is a platform for spreading information, ideas, and influence among its members [16]. Social network analysis (SNA) is one of mathematical and graphical analysis methods to represent the descriptions of networks compactly and systematically [13].

Kiss and Bichler [18] compared plenty measures of influence including different centrality measures in customer networks. They indicated that the eigenvector centrality is one of effective measures for estimating the influence of a node in a network. Hanneman [13] indicated that $m$-step reach centrality is used to measure reach efficiency (e.g. what portion of all others ego can reach in a network) in $m$ steps. On the hand, it indicates the smallest number of edges that one needs to pass through in order to reach from a given node to all the other nodes in the network [27].

In this work, SNA was used as a tool for constructing interaction network to estimate the possible information transition states. Meanwhile, it was also used to determine customer value in order to estimate the influenceability and reachability representing the possible diffusion reward of different strategies.

3. The System Framework

In this section, we proposed an Advertisement Path Planning Mechanism (APPM) to support marketers’ online information diffusion process in micro-blogosphere. The APPM treats the diffusion process as sequential optimization problem rather than a static prediction problem.

The procedures for a marketer to solicit information diffusion to his/her friend network in the context of social media marketing are described as follows.
marketer would like to propagate marketing information starting from a list of influencers/endorsers, which were previously selected or recommended by other mechanisms. The marketer might want to know the best diffusion path and the possible reward (influenceability and reachability) for different purposes, e.g., for seeking business opportunities or for building brand impression. And the starting node might want to know whom they should pass the information to within their own friend network. However, nodes might or might not pass information on the planned path; if there is one node breaking the planned path (does not pass the information to the next node as expected), the proposed system will re-plan path from the breaking point.

For example, as shown in Figure 1, a path for advertisement propagation is planned as $U_A \rightarrow U_B \rightarrow U_C \rightarrow U_D$. We could easily obtain that $U_B$ and $U_C$ respectively contained in the propagated list of $U_A$ and $U_D$, which means this path is continuing. But, obviously, $U_D$ is not contained in the propagated list of $U_C$, which means that the path is broken so that the system needs to re-plan the diffusion path for $U_I$, $U_J$, and $U_K$, which are selected as the restarting nodes to continue the information diffusion and report to marketer.

![Figure 1. Information diffusion path.](image)

Figure 2 depicts the framework of our system. The proposed framework is comprised of three main components: transition flow inferring module, customer value analyzing module, and diffusion path estimating module:

1. **Transition Flow Inferring Module**: The purpose of this module is to determine the dynamic social network factor: Interaction. The basic concept of Markov model is to determine the transition probability of transitions from one state to another. In the interaction network, the transition probabilities between possible states are estimated according to social interactions.

   **3.1. Interaction Network Construction**: We leveraged the active social interaction data from online social networks to obtain active social nodes as the possible transition states. When the circle of people’s friendship grows, there is an increasing need for friend management. Research by Dunbar [10] indicated that there is an approximate natural group size in which everyone can really know each other.

2. **Customer Value Analyzing Module**: The aim of this module is to determine the diffusion value of the nodes which are included in interaction network. The directed interaction relations are transformed into an adjacency matrix and the social factors in both the influenceability and the reachability are considered to derive the values of information diffusion of social nodes.

3. **Diffusion Path Estimating Module**: The objective of this module is to plan the diffusion path of starting node which was previously recommended by the other influencer discovering mechanisms [7, 16, 36]. A simple probability model consolidating the transition probabilities between social nodes and the values of information diffusion of social nodes is utilized to calculate the expected value of path planning.

![Figure 2. The framework of the advertisement path planning mechanism.](image)
Though one can have hundreds of online friends, most of them are just a name in his/her friend list but does not incur any social interaction. A recent study also showed that social media users have a very small number of off-line friends compared to the number of on-line friends they declare [14]. Here, the concept of two-mode social network analysis was used to construct interaction network to filter out the active friends for identifying the possible information transition states.

For constructing the directed interaction network of a specific user, we have to collect social interaction data from his/her micro-blogging sphere and define the direction of interaction flow. While a user posts a micro-blogging message, it is likely to say that he/she is expecting some responses. In the current paper, we defined the micro-blogging message poster and replier as “interaction requester” and “interaction provider” respectively. For example, $U_a$ and $U_b$ post a message in micro-blogging sphere, which means $U_a$ and $U_b$ are interaction requesters. And $U_c$ replied both of them, implying that $U_c$ is an interaction provider. Consequently, there would be “interaction” flowed from $U_c$ to $U_a$ and $U_b$ (see Figure 3 for demonstration).

3.1.2. Transition Probability Inference. After obtaining the active social nodes (possible transition states), the following formulation was used to determine the transition probability between states.

$$
P(U_j|U_i) = \frac{[\Phi_{ij}]}{[\Phi_i]},$$  \hspace{1cm} (1)

where $[\Phi_{ij}]$ denotes the number of interaction flows from $U_i$ to $U_j$, $P(U_j|U_i)$ is the probability of $U_j$ after $U_i$, $[\Phi_i]$ denotes the total number of interaction flowing out from $U_i$. Continuing the example in Section 3.1.1 (in Figure 3), the $P(U_a|U_c)$ and $P(U_a|U_b)$ are obtained as 0.5 and 0.5 respectively. Finally, the interaction network is represented as a transition matrix ($TM$):

$$TM = \left[ P(U_j|U_i) \right]_{m \times m},$$  \hspace{1cm} (2)

where $m$ denotes the total number of transition states.

![Figure 3. Directed interaction network.](image)

3.2. Customer Value Analyzing Module

The purpose of this module is to determine the static social network structure based factors: influenceability and reachability. In this module, the friendship network constructed by friend list on micro-blogosphere was used to obtain the eigenvector centrality and reach centrality for influenceability and reachability respectively.

At first, the friend network would be represented as a bipartite graph $G = (V, E)$, where $V$ denotes the vertex in the network and $E$ denotes the edges between $V$. Next, for the influenceability and reachability analysis, $G$ would be transformed to an adjacency matrix $A = (a_{ij})$, if vertex $v$ and vertex $t$ is connected, $a_{ij}=1$, otherwise $a_{ij}=0$. For convenient, we use UCINET to compute following two centrality.

3.2.1. Influenceability Analysis. In the current research, the eigenvector centrality was used to compute the influenceability of $U_i$. Conceptually, different neighbors may have different value contributing to one’s eigenvector centrality. That is, the influenceability of a $U_i$ is contributed by the influenceability of the connected neighbors of $U_i$. The influenceability of $U_i$ ($IA(U_i)$) is determined as follow:

$$IA(U_i) = \sum_{j \in SN_i} a_{ij} \times IA(U_j) / \lambda,$$  \hspace{1cm} (3)

where $SN_i$ denotes the social network of $U_i$ and $\lambda$ denotes the eigenvalue of $A$.

3.2.2. Reachability Analysis. In the current research, the averaged $m$-step reach centrality was used to evaluate the reachability of $U_i$. That is, reachability indicates how many users $U_i$ could reach in average. The reachability of $U_i$ ($RA(U_i)$) is determined as follow:

$$RA(U_i) = \sum_{m=1}^{\infty} A^m_{i,j} / m,$$  \hspace{1cm} (4)
where $m$ denotes the number of steps. The value of $m$ could be set according to the needs of marketers. According to the small world effect [28], the value of $n$ is no need to greater than 6.

3.3. Diffusion Path Estimating Module

3.3.1. Sharing Behavior Analysis. The expected value of diffusion reward is impacted by the willing-to-share of social nodes. Although a node got higher influenceability and larger reachability than others, he/she just likes to make daily chat and specific conversation with someone but do not like to share information in micro-blogsphere. If the diffusion path plans to pass through him/her, it will be easily interrupted. Due to the small character limit (140 characters) in micro-blogsphere, a URL is frequently used to make information sharing behavior. On the other hand, a message is external information sharing from other source, if it contains URL in a micro-blogging message. The following formulation is used to determine the probability of willing-to-share behavior of a social node.

$$P(wts_{t_j}) = \frac{|M_{sop}|}{|M_{post}| + |M_{rep}|}$$  \hspace{1cm} (5)

where $P(wts)$ denotes the probability of willing-to-share of the $U_j$, $|M_{post}|$ and $|M_{rep}|$ denote the total number of messages posted and the total number of message relied to others by $U_j$ respectively. $|M_{sop}|$ denotes the total number of messages containing at least an URL in $|M_{post}|$ and $|M_{rep}|$.

3.3.2. Diffusion Path Analysis. In our proposed APPM, we have combined the probability of states transition, will-to-share, and diffusion reward function as treatments to explore the diffusion path with highest reward. At first, we defined the diffusion reward function as follow:

$$R(U_j) = \alpha \times IA(U_j) + (1 - \alpha) \times RA(U_j)$$  \hspace{1cm} (6)

where $R(U_j)$ denotes the diffusion reward of $U_j$, $\alpha$ is the information diffusion strategy weighting. The optimal path planning is defined as follow:

$$\max_{i \in SN} P(U_i, U_j)^n \times P(wts) \times R(U_j)$$  \hspace{1cm} (7)

where $n$ denotes the $n$-th step in the diffusion path. In the reward function, $P(U_i, U_j)$ is treated as a decay factor for the purpose of convergency. Within a path, higher $P(U_i, U_j)$ would make the convergence speed relatively slow, which means the information could propagate far away from starting node.

4. Experiment Design

In this section, we applied the proposed mechanism in micro-blogging system to examine its effectiveness. Micro-blogging service is one of the top tools for social media marketing. We used Plurk, one of the most popular micro-blogging services, as the platform for conducting experiments. Currently, the Plurk is very popular in Asia and United States [3]. It allows users to send and response messages in short sentences (140 characters). Besides, it also attracts users to communicate each other and to share external information by embedded URL [3]. Because Plurk is popular and predominantly used for communicating and sharing, we believe that it is an excellent platform for information diffusion while marketers make social media marketing.

While studying the issues related to social networks, the snowball sampling method is very feasible for using to construct experiments [2], and it was used in our experiments. At first, we randomly select 13 active Plurk users and invite them to join our experiment. Then, each of them is asked to invite three of their friends for recruiting our participants. This procedure is repeated 3 times in order to complete the sampling process. After filtering out those users who were not willing to join our experiment and those users who is duplicated invited, there are 97 active Plurk users left and they are also the candidates of the start point of a path in the experiment.

First, for the purpose of constructing the interaction network for obtaining the transition probability, we collected the last 6 months micro-blogging message (including post and response data) from participants’ public Plurk interface. Second, with the purpose of constructing the friendship network for obtaining the influenceability and reachability, we recursively expanded friendship from participants’ friend list. Finally, there were 2,317 social nodes in the friendship network.

In the experiment, we totally diffused 40 marketing messages via two different marketing strategies: seeking business opportunities and establishing brand expression. According to previous studies, coupons promotions could cause an increase in product sales [4] and the product reviews from third-party might spread good news/impressions of brands so that it can increase the effectiveness of firms’ advertising [5]. There were totally 20 product deals/coupons advertisements for seeking business opportunities and 20 product evaluation review articles for establishing brand.
expression. The former marketing information is collected from Yahoo! Shopping, which is one of the largest online shopping sites and the latter is collected from Epinions, which is one of the most professional and famous product review platforms allowing users to share their product experiences and opinions. The advertisements were delivered with an online 5-star rating questionnaire for the marketing information receivers to feedback their acceptance and path tracking (Which friends the marketing information forward to?).

We evaluated our proposed mechanism with four typical benchmarks: (1) random advertising without our path planning mechanism (Random), (2) random advertising with our path planning mechanism (Random+Path), (3) influencer advertising without our path planning mechanism (Influencer), and (4) influencer advertising with our path planning mechanism (Influencer+Path). The random advertising method is to randomly select participates whose $P(wts) > 0$ as the starting point of information diffusion. Besides, according to Kiss and Bichler [18], out-degree centrality has a better performance in influencer identification so that we used out-degree influencer selection to select the starting point of information diffusion. There are 5 participants selected by the random selection method as the starting points of “Random” and “Random+Path” benchmarks. There are 5 participants selected by influencer identification as starting points of “Influencer” and “Influencer+Path” benchmarks. Then, we would pass non-duplicate advertisements to them (totally 20 advertisements for each diffusion strategy) for starting our diffusion experiments.

5. Experiment Results

In order to evaluate the performance of different advertising methods, we use the click-through rate (CTR) of advertisement and the receivers’ 5-star acceptancy rating feedback of received marketing message as the evaluation indicators. The former is a popularly practical indicator about advertising efficiency; the latter could evaluate the users’ impression of the received marketing message.

Intuitively, for the purpose of seeking business opportunities, it is expected to seeking the potential customer who with highly (≥ 4-star) acceptance of the product advertisement and for establishing brand expression, it is expected to seeking the potential customer who with not lowest (≥ 2-star) acceptance of the product advertisement. We compared performance using CTR with different star rating conditions.

5.1. Seeking Business Opportunities Strategy

Generally, business opportunities exist in the potential customers who with high acceptance of product advertisement which means that he/she have a chance to buy products. The CTR with acceptance condition formula is defined as:

$$CTR = \frac{\left| \Phi_{\text{click}} \cap \Phi_{4-\text{star}} \right|}{\left| \Phi_{\text{ad}} \right|}$$

where $\left| \Phi_{\text{ad}} \right|$ denotes the total number of delivered advertisements, $\Phi_{\text{click}}$ denotes the set of clicked/read advertisements, and $\Phi_{4-\text{star}}$ denotes of advertisements which have receiver rating ≥ 4-stars acceptance.

The Figure 4 shows the CTR in different benchmark methods. The advertisements in “Random” and “Random+Path” totally respectively diffused 153 and 276 times and got 0.111 and 0.210 CTR which means that our path planning mechanism improved approximately 10% chance for seeking business opportunities. The advertisements in “Influencer” and “Influencer+Path” totally respectively diffused 392 and 517 times and got 0.298 and 0.366 CTR which means that our path planning mechanism improved approximately 7% chance for seeking business opportunities.

5.2. Establishing Brand Expression Strategy

The purpose of this marketing strategy is to enhance (4-5 stars) or reverse (2-3 starts) brand expression of customers. However, it is very hard to reverse brand expression of antis (0-1 star). Even, it might opposite effect in marketing. The CTR with acceptance condition formula is defined as:

$$CTR = \frac{\left| \Phi_{\text{click}} \cap \Phi_{2-\text{star}} \right|}{\left| \Phi_{\text{ad}} \right|}$$

where $\left| \Phi_{\text{ad}} \right|$ denotes the total number of delivered advertisements, $\Phi_{\text{click}}$ denotes the set of clicked/read
advertisements, and $\Phi_{2\text{-star}}$ denotes of advertisements which have receiver rating $\geq 2$-stars acceptance.

The Figure 5 shows the CTR in different benchmark methods. The advertisements in “Random” and “Random+Path” totally diffused 285 and 443 times and got 0.284 and 0.404 CTR, which means that our path planning mechanism improved approximately 15% chance for establishing brand expression. The advertisements in “Influencer” and “Influencer+Path” totally respectively diffused 601 and 887 times and got 0.522 and 0.592 CTR, which means that our path planning mechanism improved approximately 7% chance for establishing brand expression.

$$EA = \frac{|\Phi_{\text{receivers}}|}{|\Phi_{\text{mi}}|},$$

where $|\Phi_{\text{receivers}}|$ is the total number of receivers in addition to the path nodes and $|\Phi_{\text{mi}}|$ denotes the total number of delivered marketing information. EA is the average receivers per marketing information.

From Figures 6 and 7, we observe that the proposed APPM could enhance the exposure ability of product advertisements, if we ignore the acceptance of product advertisement. For the random advertising method, APPM respectively improved approximately 80% and 60% exposure ability of the random advertising method in the seeking business opportunities strategy and in the establishing brand expression strategy. For the influencer advertising method, APPM respectively improved approximately 32% and 50% exposure ability of the random advertising method in the seeking business opportunities strategy and in the establishing brand expression strategy.

5.3. Exposure Ability in Different Strategies

Advertisers concern about the effective exposure for their advertisements. The proposed APPM would plan suitable path for advertisements in different strategies. In one of diffusions, the total number of message receivers in addition to the people who are included in the planned diffusion path is the message exposure range of path planning. For instance, as Figure 1, the $U_E, U_F, U_G$, and $U_H$ are the exposure range of the $U_A \rightarrow U_B \rightarrow U_C \rightarrow U_D$ that is a planned diffusion path. Because the path was broken by $U_D$ and we re-plan diffusion path for $U_I, U_J$, and $U_K$, the planned diffusion path of the diffusion would be adjusted as

$$U_A \rightarrow U_B \rightarrow U_C \rightarrow U_I \rightarrow \cdots \rightarrow U_J \rightarrow \cdots \rightarrow U_K \rightarrow \cdots.$$

However, the re-planned diffusion paths still belong to the same marketing information diffusion. The eventual number of message receivers of diffusion is an important indicator to evaluate the performance of planned diffusion path. The exposure ability (EA) is formulated as follows:

$$EA = \frac{|\Phi_{\text{receivers}}|}{|\Phi_{\text{mi}}|},$$

6. Discussion and Conclusion

In this paper, we proposed an advertisement path planning mechanism, named APPM, which is based on probability and optimization models. Our mechanism
treated the diffusion problem as a sequential optimization problem. It can support the marketers to evaluate the possible information diffusion effectiveness under different marketing strategies and support the evaluated influencers propagating information to specific individuals for continuing the diffusion process. We adopted transition flow inferring, customer value analyzing, and diffusion path estimating modules to plan the optimal diffusion path for influential social nodes.

Our experiment results got a positive outcome. Performance evaluations validate that the proposed mechanism can significantly improve the diffusion process of advertising messages and decrease the marketing uncertainty of marketers while they decide to deliver information for social media marketing. Even using random influencer selection for selecting diffusion starting points, the proposed path planning mechanism could support and improve the diffusion effectiveness approximate to influencer advertising method. And, this mechanism would be able to play a greater performance, if combined with other influencer discover mechanisms.

There are still some aspects that can be further improved. While determining the possible transition states and diffusion reward, the preference and reputation of social node should be taken into account. The preference similarity might be used to evaluate the fitness between types of advertisement and preference of user. These two factors might be positive factors related to reward function. Besides, the optimal path planning formulation might be subject to some conditions, for example, both of influenceability and reachability of social node should greater than a threshold.

7. Acknowledgements

The authors are grateful for the financial support of National Science Council (NSC: 101-2410-H-009-009-MY3).

8. References


