

## Accepted Manuscript

Title: A Diffusion Planning Mechanism for Social Marketing

Author: Yung-Ming Li Cheng-Yang Lai Lien-Fa Lin

PII: S0378-7206(16)30415-3

DOI: <http://dx.doi.org/doi:10.1016/j.im.2016.12.006>

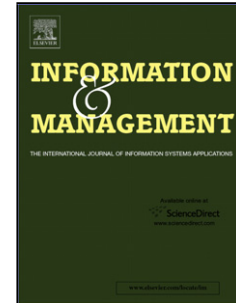
Reference: INFMAN 2964

To appear in: *INFMAN*

Received date: 16-10-2015

Revised date: 31-10-2016

Accepted date: 18-12-2016



Please cite this article as: Yung-Ming Li, Cheng-Yang Lai, Lien-Fa Lin, A Diffusion Planning Mechanism for Social Marketing, Information and Management <http://dx.doi.org/10.1016/j.im.2016.12.006>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## **A Diffusion Planning Mechanism for Social Marketing**

Yung-Ming Li<sup>a</sup>, Cheng-Yang Lai<sup>a</sup>, Lien-Fa Lin<sup>b</sup>

<sup>a</sup>Institute of Information Management, National Chiao Tung University, Hsinchu, 300, Taiwan

<sup>b</sup>Department of Information Communication, Kao Yuan University, Kaohsiung, 821, Taiwan

\*Corresponding author: Yung-Ming Li (yml@mail.nctu.edu.tw)

E-mail: Cheng-Yang Lai (cylai.k@gmail.com), Lien-Fa Lin (lienfa0704@gmail.com)

## Abstract

Social media is gaining importance as a component of marketing strategies. Many types of social media, such as social networking sites, blogospheres and micro-blogospheres, have been seeking business opportunities and establishing brand expression in the recent years. Online marketing information diffusion has become the critical business model of online social networks. However, most of the current marketing studies on discovering potential influences do not appropriately support them to diffuse advertisements. Therefore, marketing information may be lost during the diffusion process and cannot be sent to potential customers successfully. In this research, a diffusion path planning mechanism for advertisements is developed to help influencers to propagate marketing information and help marketers to evaluate possible rewards under different marketing strategies. Our experimental results show that the proposed mechanism can significantly improve the diffusion process of advertising messages and decrease marketing uncertainty.

**Keywords:** information diffusion, social network, social media marketing, path planning, influential nodes

## 1. Introduction

### 1.1. Background

Traditional media is a one-way communication system—a brand generates a message and transmits it to the masses through broadcasting, print, radio or television, which have their limitations. According to the Nielsen report, 92% of consumers believe word of mouth more than traditional advertising. Moreover, traditional advertising techniques cannot be implemented easily because they do not provide a platform for more users to share their personal profile with people they do not know, and therefore, it is difficult to target certain types of users using traditional filtering techniques. Marketers know word-of-mouth messages are much more powerful than brand messages. However, they do not really provide the tools to generate word-of-mouth diffusion in traditional media.

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), has provided powerful means of organizing friend networks, publishing contents and sharing information [43]. The importance and polarity of social media are continually increasing in people's daily life. Nowadays, most online users will join in at least one form of social media to obtain and share information not only for personal use but also for updated information about a company, brand or product [38]. The growing user population of social media reveals the importance of social media in business usage, particularly in the marketing field, e.g. viral marketing and online advertising.

Advertising through online media has greatly increased year after year (22% increase from 2010 to 2011) [4]. According to the reports by Stelzner [51] and Nielsen [45], 83% of marketers said that social media was important to their promotional business strategies, and 93% of companies always use social media for marketing purposes (50% of these companies had applied social media in their marketing strategies for more than 1 year, and 73% of these companies planned to increase the marketing uses of social media). By exploiting social relation analysis and social activities mining, we can obtain more useful information from social media. Obviously, companies will lose lots of opportunities to reach potential consumers if they are not engaged in social networking media as a part of their online marketing strategy [55].

Social marketing, i.e. delivering of marketing information through social media, has become one of the most significant promotion methods for businesses. It can exploit the power of social influence and word of mouth to deliver marketing information. Even more, it might lead to the information cascade effect [23] by stimulating the buying intentions and turning the brand impressions of potential consumers. Recently, sellers (enterprises and individuals) have promisingly turned to propagating marketing information through online media for seeking business opportunities (e.g. product advertisements) [39, 53, 54] and establishing brand expression (e.g. branding messages) [33, 37, 38].

Information diffusion through online social networks has recently become an active research topic [2]. According to Brown and Hayes [12], the crucial work of influencer marketing is to identify the influencer or endorser for diffusing information, known as the ‘key player problem’ (KPP) [10]. In addition, a way to help the identified key player to disseminate information is also needed. Generally, to the best of our best knowledge, information diffusion-related research applies related analyzing techniques (e.g. social network analysis) to identify the powerful influencers or endorsers who might most help to diffuse [1,19,57]. However, how to propose the appropriate diffusion path planning to help them to deliver marketing information to obtain better marketing effectiveness (e.g. raising product sales or gaining brand awareness) has rarely been studied. Influencers and endorsers are commonly selected through recommender systems, which are expected to reach and influence potential customers [7, 35, 40]. However, it is not well known how to guide and support these influencers/endorsers regarding passing the marketing information on. In other words, which direction is best for the diffusion process if the information to be diffused starts from him/her while an evaluated influencer/endorser receives the marketing information?

## **1.2. Research Problems and Methods**

Because Internet is now the primary message-delivering medium, it is worthwhile to investigate and design a novel mechanism for supporting online marketing information propagation. Although marketers include online media as part of marketing strategies, the measurement of social media effects is an increasing concern [27]. Specifically, the aim of our research is to identify 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 to gain the required diffusion reward (e.g. information influenceability and target reachability).

To address these issues, in this research, we design a diffusion path planning mechanism to help influencers/endorsers to diffuse information. The proposed mechanism considers the factors of user preferences, social interaction transition probabilities, network influenceability and reachability, and willingness to share, increase the effectiveness of advertisements and support the selected influencers/endorsers in delivering information to appropriate friends. The aim of the proposed mechanism is to identify the key players among the selected influencers/endorsers to continue to boost their social influenceability and reachability for maximizing the advertising effectiveness for business. We validate our mechanism by conducting experiments in Plurk, one of the most popular micro-blog platforms. Our experimental results show that the proposed model could enhance the advertising effectiveness of influencer marketing.

The remaining paper is organized as follows. The basic concepts and literature related to our research topics are provided in Section 2. In Section 3, we present the proposed social diffusion support framework, which integrates the techniques of latent semantic indexing (LSI)-based preference analysis, social network analysis, Markov chain propagation transition and optimal diffusion path planning. Practical experiments are described in Section 4, and the results of the experiment are discussed in Section 5. Finally, in Section 6, some research contributions and several directions for future studies are provided and discussed.

## **2. Literature Review**

In this section, we will review the concepts of social media marketing, information propagation theories and theories of dynamic social impact, key influencer selection, and path planning for marketing information propagation techniques.

## 2.1. Social Media Marketing

Social media marketing is a new and rapidly growing way in which businesses are reaching out to potential customers. It refers to the process of gaining users' attention and acceptance through social media. Social media such as Facebook, Plurk, and Twitter are online platforms that are used to deliver information through social interactions (e.g. communication with family, colleagues, and friends) [48]. Jackson [31] has shown that online media are more effective in influencing consumers than classic marketing channels. The most common theories within the domain of social marketing are social norms theory (deciding which messages are appropriate and relevant for which audience salience/creating credible messages in terms of message, source and explanation of data), the health belief model, the theory of reasoned action/theory of planned behaviour, social cognitive theory, the transtheoretical model of health behaviour change and diffusion of innovations. Because consumers have begun to mistrust and refuse to accept official advertising [59], a message would be more acceptable if it were delivered from their close friends [7, 59]. 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 sites, blogospheres and micro-blogospheres [42]. For example, Iyer et al. [30] examined the advertising strategy and found that firms' advertising strategy should focus more on those consumers who have a strong preference for their product. Yang et al. [58] proposed a data mining framework on the basis of the customer's interaction data from social networks to support online advertising. Kazienko and Adamski [32] proposed the AdROSA system for personalized web advertising, which integrates web usage and data mining techniques to reduce user input and to respect users' privacy. Social media marketing has become such an important feature that it is no longer a question of whether to use it but how to use it [44,61]. Therefore, in this study, we propose a mechanism for planning the diffusion path to support social media marketing.

## 2.2. Information Propagation

Information propagation in online social networking sites has attracted great research interest recently. Informative and persuasive diffusion are the two main purposes of the information diffusion

process [5,7,32,56]. Informative diffusion focuses on delivering information to receivers who are extremely interested in it. In the marketing field, for example, marketers could perform informative diffusion to deliver promotional information about products to consumers seeking business opportunities. Persuasive diffusion focuses on delivering information to impress the receivers. In the marketing field, for example, marketers could perform persuasive diffusion to deliver branding information about products to consumers for establishing brand impressions.

Information diffusion techniques in social networks are broadly used for influencing and informing people [46]. According to the social impact theory [60], increasing the strength of a source of influence increases the influence on the target. Credibility refers to a person's perception of the truth of a piece of information. Source credibility theory has been proposed in WOM communication studies of consumer psychology and marketing [22]. The positive effects of viral marketing to influence [7] and word of mouth [25] to inform potential consumers have been observed. Obviously, information (e.g. informative and persuasive information) provided by friends is more trustable and acceptable than that from marketers [34,59]. Peer influence means that an individual might lead other individuals to do things according to the information gathered from him/her. Park and Kim [47] focussed on revealing the effect of persuasive information (online consumer reviews) on purchasing intention for experts and novices. Abubakre et al [1] and Yang and Leskovec [57] focussed on effective ways of diffusing informative promotional information about products. However, marketers would not focus on just one strategy for marketing. This research proposes a hybrid marketing strategy that considers both purposes of raising product sales and brand awareness.

### **2.3. Key Player Problem**

The KPP is a procedure that finds a set of key players in a social network for different purposes. Borgatti [10] defined KPP positive (KPP-POS) and KPP negative (KPP-NEG) as two related problems for discovering sets of key players. KPP-POS is defined as the identification of key players who could be used as seeds for diffusing some information on the network. KPP-NEG is defined as the identification of key players who could be used as the breaking points for disrupting or fragmenting the network. However, the research field of social media marketing mainly focuses on



KPP-POS for maximizing advertising effectiveness.

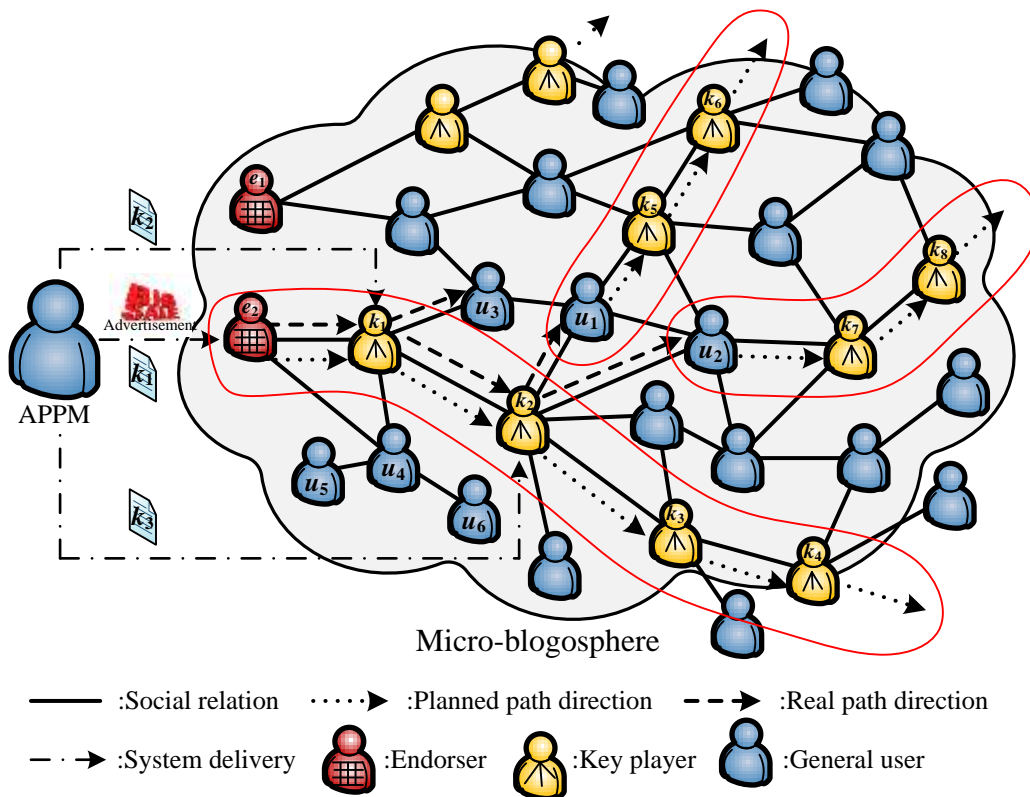
Previous works have shown that peer influence has a positive effect in online marketing [14,15,18,42,52] in selecting the key player for marketing purposes. As a result, influence-quantifying models have been proposed for solving the KPP-POS problem. Yang and Leskovec [57] developed a linear influence model to predict the possible influential nodes in the network for modelling the information diffusion in online social media. Lin et al. [41] proposed an endorser-based social diffusing mechanism, which considers the factors of preferences, i.e. influence and the diffusion power of users, to enhance the effectiveness of target advertising by discovering the most appropriate endorsers that can propagate the ads to the identified target users. This work focuses on choosing initial seed endorsers. It discovers potential influencers (endorsers) but does not appropriately help them to diffuse advertisements by continuous diffusion path planning. However, according to Brown and Hayes [12], implementing influencer marketing not only begins with the key influencer selection but also looks for a way to work with them to help them do their job better.

This research proposes a mechanism that could sequentially select the key players from among the identified key influencers to continue to boost their advertising effectiveness. In this research, the information diffusion problem is seen as a sequential path planning optimization problem rather than a simple recommender problem. Therefore, we propose a simple maximization model for planning the optimal diffusion path for influencers who are evaluated by other mechanisms.

### **3. The System Framework**

In this section, we propose an advertisement path planning mechanism (APPM) to support marketers' online information diffusion process in micro-blogspheres. The APPM plans the diffusion process as a dynamic path optimization problem rather than a static node influence prediction problem. The procedures for conducting information diffusion through social media marketing are described as follows. A marketer propagates marketing information by distributing the ads to the starting endorsers, which can be selected according to some evaluation criteria such as influence or active strength. For each starting node, we recommend the diffusion path that is generated depending on the aggregate reward, which is measured by information influenceability and ad reachability. In the

mechanism, a diffusion path is generated for aggregated reward maximization. A starting node is only aware of the first node in the planned diffusion path and decides whether to forward the ad to the node spontaneously. If a node breaks the planned diffusion path (does not pass the marketing information to the next node as planned in the diffusion path), the proposed system will replan a diffusion path from the breaking node. For example, as shown in Figure 1,  $e_2$  is one of the marketer-identified starting endorsers and a diffusion path for marketing information propagation is planned as  $e_2 \rightarrow k_1 \rightarrow k_2 \rightarrow k_3 \rightarrow k_4$ . The system will first deliver the information to  $e_2$  and suggest the next key player ( $k_1$ ) to forward to. If  $k_1$  receives the information, then the system will suggest that he/she should forward the information to  $k_2$  and so on. If  $k_2$  receives the information but  $k_2$  breaks the planned diffusion path (i.e. does not forward to the suggested  $k_2$ ) and passes the marketing information to  $u_1$  and  $u_2$  so that APPM re-plans the diffusion path  $u_1 \rightarrow k_5 \rightarrow k_6$  and  $u_2 \rightarrow k_7 \rightarrow k_8$  for  $u_1$  and  $u_2$ , respectively, to continue the marketing information diffusion process.



**Figure 1.** Information diffusion path.

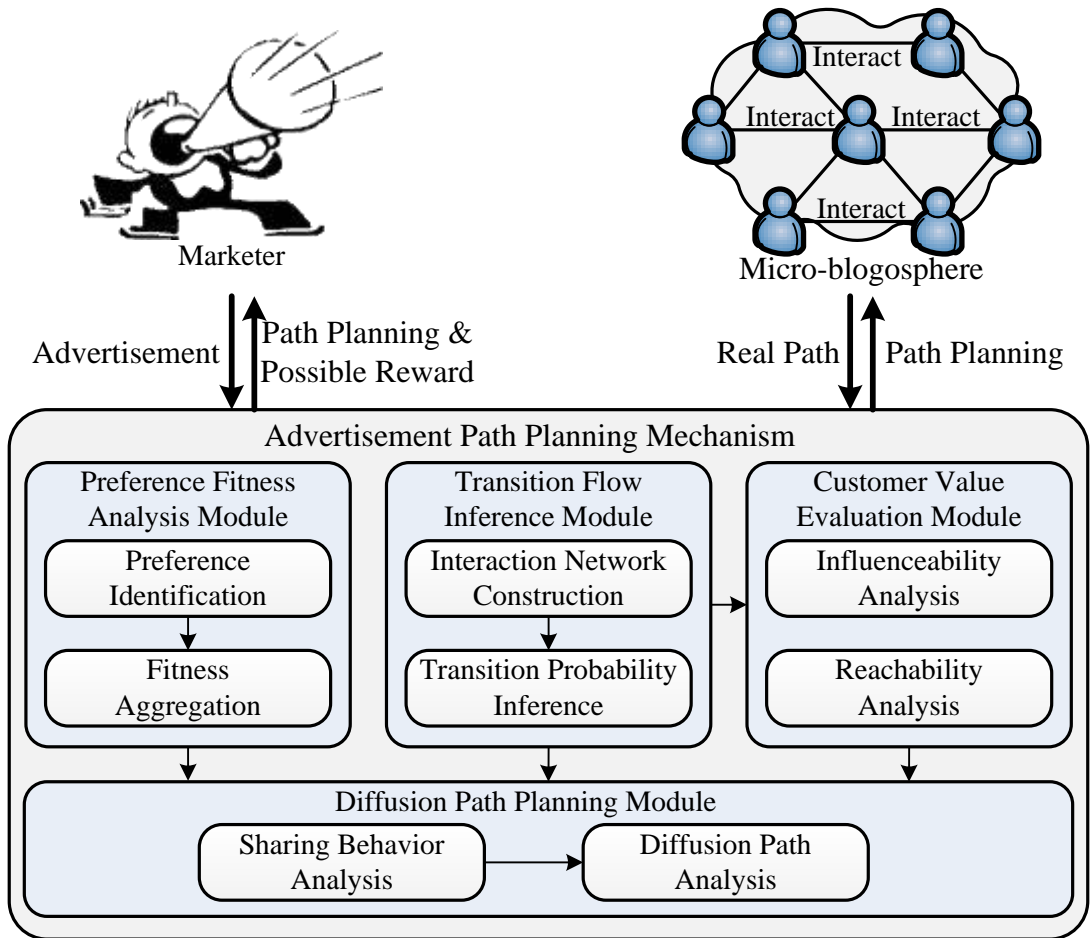
Figure 2 shows the framework of the proposed system. The proposed framework comprises four main components: the preference fitness analysis module, transition flow inferring module, customer value analyzing module and diffusion path planning module.

(1) Preference fitness analysis module: preference fitness analysis is used to measure the fitness degree between a user's preference and marketing information. The LSI-based methodology is exploited to estimate the preference fitness of a user for marketing information by analyzing their daily micro-blogging messages.

(2) Transition flow inferring module: the purpose of transition flow inferring analysis is to infer the transition probability of the possible information forwarding between two users according to the daily social interactions among the users within the social network. We apply the Markov chain concept to derive the transition probabilities of information forwarding.

(3) Customer value evaluation module: the aim of customer value evaluation is to estimate the diffusion value of the nodes that are included in the social network according to the interaction intensity. The directed interaction relations are transformed into an adjacency matrix, and the diffusion effectiveness factors, namely influenceability and reachability, are considered to produce the information diffusion value of a node.

(4) Diffusion path planning module: the objective of diffusion path planning is to identify the optimal diffusion path starting from a seed endorser node, which could be recommended by the influencer discovering mechanisms [1,19,57]. The path that maximizes the aggregate diffusion reward is generated by integrating the propagation tendency (transition probabilities between social nodes) and propagation reward (information diffusion value of social nodes).



**Figure 2.** The framework of the APPM.

### 3.1. Preference Fitness Analysis Module

As users have a higher tendency to share with friends the information they are interested in, it is essential to analyze the matching between the preference of a user and the information to diffuse. The preferences of users could be discovered according to the information that they shared on social media. For example, a preference of a user would be represented by the micro-blogging messages he/she had posted on the micro-blogsphere.

#### 3.1.1. Preference Identification

In this research, the LSI technique [20] is used to model the user's preference for specific products. LSI, one of the well-known information retrieval algorithms, is a process used for mapping keywords to a vector and finding the most relevant documents from a group of documents. In practice, the marketer could give some keywords that most represent their products to promote them and users' preferences can be implicitly discovered from their daily sharing behaviours; therefore, LSI

would be an appropriate method for this research for identifying preferences. The procedures of LSI-based preference identification are described as follows:

Step 1. Construct the term–post matrix and keywords of the product column matrix.

For each user, the micro-blogging messages posted in the last 6 months are gathered to represent his/her preference. Then, each post included in the corpus is tokenized, and the stop words in the post are removed for extracting the terms. A term–post matrix (TD) that consists of  $m$  terms and  $n$  posts can be expressed as

$$TD = [tf_{ij}]_{m \times n},$$

where  $tf_{ij}$  denotes the term frequency of term  $i$  in post  $j$  of the corpus, and it is simply defined as the total occurrence of term  $i$  in post  $j$ .

For estimating the LSI-based product–user similarity, the representative keywords for the product are required and could be given by the marketer. The product keyword column matrix (KC) can be expressed as

$$KC = [ko_{ij}]_{m \times 1},$$

where  $ko_{ij}$  denotes the occurrence of keyword  $i$  in term  $j$ . If keyword  $i$  hits term  $j$ ,  $ko_{ij} = 1$ , otherwise  $ko_{ij} = 0$ .

For example, a user posted three micro-blogging messages and the product keywords given by the marketer are as follows:

p1: *Wow~ It's really sunny today~ Summer is coming~!*

p2: *Sunburned! I would use the high SPF sunblock lotion and I would not get sunburned again.*

p3: *I've been looking for good sunscreen that will work even while I'm sweating.*

keywords: *sunblock, sunscreen, lotion, sunburned, SPF*

Matrices  $TD$  and  $KC$  can be constructed and represented as

<i>terms</i>	<i>p1</i>	<i>p2</i>	<i>p3</i>	<i>keywords</i>
<i>good</i>	0	0	1	0
<i>high</i>	0	1	0	0
<i>lotion</i>	0	1	0	1
<i>spf</i>	0	1	0	1
<i>summer</i>	1	0	0	0
<i>sunblock</i>	0	1	0	1
<i>sunburned</i>	0	2	0	1
<i>sunny</i>	1	0	0	0
<i>sunscreen</i>	0	0	1	1
<i>sweating</i>	0	0	1	0
<i>today</i>	1	0	0	0
<i>work</i>	0	0	1	0
<i>wow</i>	1	0	0	0

Step 2. Compute the similarity between micro-blogging message and keywords of product.

In this step, the singular value decomposition (SVD) [13,24], a well-known matrix factorization technique, is used to decompose the  $TD$  into three matrices. Because the SVD method provides the lower-rank approximations of the matrix, it is very useful for our application. SVD can produce a low-dimensional representation of the  $TD$ , and the original matrix can be obtained through the following matrix multiplication:

$$TD = U \cdot \Sigma \cdot V^T,$$

where matrices  $U$  and  $V^T$  are two orthogonal matrices and  $\Sigma$  is a diagonal matrix with all singular values of the matrix  $TD$  as its diagonal entries. All the entries of matrix  $\Sigma$  are positive and stored in the decreasing order of their magnitude.

LSI retains only the first  $k$  singular values together with the corresponding rows of  $U$  and  $V$ , which induce an approximation to  $TD$ . The dimensionality reduction obtained by performing SVD reduced matrix  $\Sigma$  to have only  $k$  (a tuned parameter) largest diagonal values ( $\Sigma_k$ ). Accordingly, although matrix  $U$  and matrix  $V$  are both reduced, the reconstructed matrix  $TD_k = U_k \cdot \Sigma_k \cdot V_k^T$  is the closest rank- $k$  matrix to  $TD$ . In other words, the dimensionality reduction in the SVD method projects large dimensions (may have thousands of dimensions) into much smaller dimensions ( $k$  dimensions). Each row of  $U_k$  represents a term as a  $k$ -dimensional vector and each row of  $V_k$

represents a post as a  $k$ -dimensional vector. From the matrix  $V_k^T$ , we can deduce that this matrix must contain  $n$  number of rows holding eigenvector values for  $n$  posts. Each of these rows then holds the coordinates of individual post vectors (PV). Each PV represents an individual post.

However, the selection of  $k$  value, which is the reduced dimensional representation, is an active open research area. It is difficult to find the best one and it is usually determined through sequential experiment tests [17]. According to previous studies [17,20,24,36], a  $k$  value of around 100 would give better performance. Therefore, the value of  $k$  was set to 100 in this study.

Step 3. Incorporate the product keywords and compute the preference similarity.

To incorporate the keywords for a product, we used the definition described by [9] to obtain the keyword vector (KV) for computing the similarity with the user's preference, and this is defined as

$$KV = KC^T \cdot U_k \cdot \Sigma_k^{-1}.$$

After obtaining KV, we computed the cosine similarity [50] between KV and each PV of users as

$$sim(KV, PV^i) = \frac{\sum_{j=1}^k (KV_j \times PV_j^i)}{\sqrt{\sum_{j=1}^k KV_j^2} \sqrt{\sum_{j=1}^k (PV_j^i)^2}},$$

Where  $PV^i$  denotes the  $i^{th}$  post vector of the user (each post vector represents a micro-blog post), and  $KV_j$  and  $PV_j^i$  indicate the  $j^{th}$  element of KV and  $PV^i$ , respectively.

### 3.1.2. Fitness Aggregation

Although user preference can be observed by their posts, the importance levels of these posts should be different when they are used for evaluating the user preference fitness to the product. For example, two articles highly correlated with the product were posted yesterday and 3 months ago. The former means the user is now focussing on the related product information, so that the user would be more willing to adopt and share the product information. However, the latter might indicate that the user once surveyed the related information but might not still be interested in the related product information if his/her focus has changed. A preference weighting function for article  $i$ , which decreases with time, is defined as

$$pw(t_i) = \frac{1}{t_i},$$

where  $t_i$  denotes the time period since the article  $i$  was posted. For example,  $t_i = 1$  indicates that the article was posted within the last month. Finally, the preference fitness of user  $U$  to the product is formulated as

$$PF(U) = \frac{\sum_{i=1}^n pw(t_i) \times \text{sim}(KV, PV^i)}{n}, \text{ where } n \text{ is the total number of articles posted by user } U.$$

### 3.2. Transition Flow Inference Module

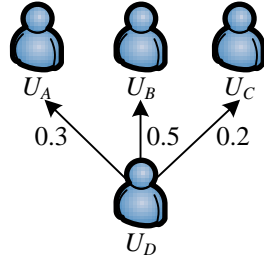
The basic concept of a Markov model is to determine the transition probability of transitions from one state to another. In the context of social networks, a state stands for a user and transition between two states is interpreted as interaction between two users. Specifically, the transition probabilities between possible states are estimated according to social interactions.

#### 3.2.1. Interaction Network Construction

We leveraged social interaction data from online social networks to obtain a set of active social nodes with respect to a specific user and used the identified nodes as the possible transition states from the current state (the specific user). When the circle of people's friendship grows, there is an increasing need for friend management. Research by Dunbar [21] indicates that there is an approximate natural group size in which everyone can really know each other. Although one can have hundreds of online friends, most of them are just a name on one's friends' list and do not incur any social interaction. A recent study also shows that social media users have a very small number of offline friends compared with the number of online friends they declare [29]. We construct a network of social interactions to filter out the active friends of a user and use these nodes as the possible information transition states. Specifically, the directed interaction network of a specific user is constructed by analyzing the social interaction data collected from his/her micro-blogsphere. The edge direction in the interaction network represents the direction of interaction flow. When a user posts a micro-blogging message, he/she is likely expecting some responses. In the current paper, we define a micro-blogging message poster and replier as 'interaction requester' and 'interaction



provider', respectively. For example, as shown in Figure 3,  $U_A$ ,  $U_B$  and  $U_C$  post messages in their micro-blogspheres, which means  $U_A$ ,  $U_B$  and  $U_C$  are interaction requesters. And  $U_D$  replies to all of them, implying that  $U_D$  is an interaction provider. Consequently, there would be 'interaction' flowing from  $U_D$  to  $U_A$ ,  $U_B$  and  $U_C$ .



**Figure 3.** Directed interaction network.

### 3.2.2. Transition Probability Inference

After obtaining the set of active social nodes (possible transition states), the following formulation is used to determine the transition probability between states:

$$P_r(\overline{U_i U_j}) = \frac{|\Phi_{\overline{U_i U_j}}|}{|\bigcup_j \Phi_{\overline{U_i U_j}}|},$$

where  $|\Phi_{\overline{U_i U_j}}|$  stands for the number of interaction flows from  $U_i$  to  $U_j$ ,  $P_r(\overline{U_i U_j})$  is the interaction transition probability from  $U_i$  to  $U_j$  and  $|\bigcup_j \Phi_{\overline{U_i U_j}}|$  denotes the total number of interactions flowing out from  $U_i$ . As shown in Figure 3,  $P_r(\overline{U_D U_A})$ ,  $P_r(\overline{U_D U_B})$  and  $P_r(\overline{U_D U_C})$  were obtained as 0.3, 0.5 and 0.2, respectively. Finally, the interaction network is represented as a transition matrix ( $TM$ ):

$$TM = \left[ P_r(\overline{U_i U_j}) \right]_{m \times m}, \text{ where } m \text{ denotes the total number of active social nodes.}$$

### 3.3. Customer Value Evaluation Module

The purpose of this module is to evaluate the network structure-based measurements: influenceability and reachability. In this module, the friendship network constructed by the friends' list on the micro-blogsphere is used to obtain the eigenvector centrality and reach centrality for

evaluating the influenceability and reachability, respectively.

At first, the friends' network would be represented as a bipartite graph  $G=(V,E)$ , where  $V$  denotes the vertices 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  $A=(a_{v,t})$ ; if vertex  $v$  and vertex  $t$  are connected,  $a_{v,t}=1$ , otherwise  $a_{v,t}=0$ . In this research, we use UCINET<sup>1</sup> to compute the following two measurements of centrality.

### 3.3.1. Influenceability Analysis

For business, the main interest of the marketers is to know how many purchase intentions of potential consumers may possibly be stimulated by the marketing information they have received. In this respect, the influence of a node plays an important role in enhancing the diffusion effectiveness of marketing information for seeking business opportunities. Kiss and Bichler [35] compared several measures of influence including different centrality measures in customer networks and suggested that the eigenvector centrality is one of the most effective measures for estimating the influence of a node in a network. In the current research, the eigenvector centrality was used to compute the influenceability of the users. Conceptually, different neighbours may have different values contributing to one's eigenvector centrality. That is, the eigenvector centrality of a user  $U_i$  is contributed by the eigenvector centrality of the connected neighbours of  $U_i$ . The eigenvector centrality of  $U_i$  is determined as

$$ec(U_i) = \frac{\sum_{j \in SN_i, j \neq i} a_{i,j} \times ec(U_j)}{\lambda},$$

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

For comparisons within a graph, it is suitable to use the eigenvector centrality with maximum normalization [49], which is derived as

$$ec_{\text{norm}}(U_i) = \frac{ec(U_i)}{\max_j ec(U_j)}.$$

---

<sup>1</sup> <https://sites.google.com/site/ucinetsoftware/home>

According to the network structure, a person with higher centrality could influence other nodes more in a social network. In addition, one person is influenced by another through the social interactions between them. Therefore, the influenceability of  $U_i$  is measured as

$$IA(U_i) = ec_{norm}(U_i) \times asn(U_i),$$

where  $asn(U_i)$  is the total number of active social nodes with respect to  $U_i$ .

### 3.3.2. Reachability Analysis

With regard to establishing brand expression, the number of potential consumers whose performance can be reached during the marketing information diffusion process is what the marketers care most about. Hanneman [26] suggested an  $m$ -step reach centrality [11] to measure the reach efficiency (e.g. the proportion of all others' ego that can be reached in a network). In the current research, the  $m$ -step reach centrality is used to evaluate the reachability of  $U_i$ . The  $m$ -step reach centrality measures the number of reachable nodes within  $m$  steps from a given social node. That is, reachability indicates how many users  $U_i$  could be reached on average per step. The reachability of  $U_i$  is measured as

$$RA(U_i) = \frac{\sum_{n=1}^m F(n, U_i)}{m},$$

where  $m$  denotes the number of steps and  $F(n, U_i)$  is the number of nodes that can be reached by  $U_i$  at  $n$  steps. The value of  $m$  could be set according to the needs of marketers. According to the small-world effect [44], the value of  $n$  does not need to be greater than 6.

## 3.4. Diffusion Path Planning Module

### 3.4.1. Sharing Behaviour Analysis

The expected value of diffusion reward is impacted by the willingness to share social nodes. Even if one node obtains higher influenceability and larger reachability than others, then the user might just like to chat daily and have a specific conversation with someone but does not like to share information in a micro-blogsphere. If the diffusion path is planned to pass through him/her, it will be easily interrupted. Because of the small character limit (140 characters) in micro-blogspheres, a URL

is frequently used to promote information sharing behaviour. However, a message is external information sharing from other sources if it contains a URL in a micro-blogging message. The degree of daily sharing behaviour of a social node is measured as

$$sb(U_i) = \frac{|M_{\text{http}}|}{|M_{\text{post}}| + |M_{\text{reply}}|},$$

where  $|M_{\text{post}}|$  and  $|M_{\text{reply}}|$  denote the total number of messages posted and the total number of messages replies to others by  $U_i$ , respectively.  $|M_{\text{http}}|$  denotes the total number of messages containing at least one URL in  $|M_{\text{post}}|$  and  $|M_{\text{reply}}|$ .

According to a previous survey [28], egoism and altruism are the two significant motivations of users who are willing to share information. Egoism refers to users who would like to share information that they have preferences about with friends because they believe their sharing behaviour could enhance their personal reputation. Altruism refers to users who are willing to enhance the welfare of their friends without expecting anything in return. Therefore, users would like to share information with friends because the users might know their friends' preferences.  $U_i$ 's willingness to share is defined as follows:

$$wts(U_i) = PF(U_i) \times sb(U_i).$$

### 3.4.2. Diffusion Path Analysis

In the proposed APPM, the probability of state transition, the tendency of willingness to share and diffusion reward function were combined as treatments to explore the diffusion path with the highest diffusion reward. First, we define the diffusion reward function as

$$DR(U_i) = \alpha \times IA(U_i) + (1 - \alpha) \times RA(U_i),$$

where  $\alpha$  is the information diffusion strategy weighting for balancing the performance of influenceability and reachability, which is determined by the focus of marketing strategies (business opportunity seeking or brand awareness). The direct reward coming from a neighbour node  $i$  to starting node  $s$  can be formulated as

$$Neighbor\_DR(s,i) = P_r(\overline{U_s U_i}) \times wts(U_i) \times DR(U_i).$$

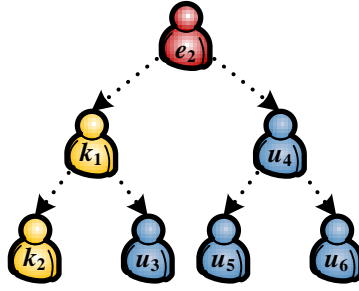
The total reward generated from diffusing the information through the planned optimal path, which is started from node  $s$ , is defined as

$$TR(s) = \max_{i \in SN, i \notin Path(s)} (Neighbor\_DR(s,i) + TR(i)), \quad Path(s) = \{s\} \cup Path(i),$$

where  $TR(i) = 0$  and  $Path(i) = \emptyset$  for  $Neighbor\_R(s,i) = 0$  or  $Path\_Length(s,i) \geq \hat{l}$ .

$Path(s)$  consists of a sequentially selected key endorser node in the social network.  $Path\_Length(s,i)$  denotes the path length between node  $s$  and node  $i$ .  $\hat{l}$  stands for the maximal length of a planned path.  $TR(s)$  is the conservatively estimated reward of the diffusion process along the path starting from node  $s$ . That is, if the marketing information could be disseminated by following  $Path(s)$ , the marketer could gain the diffusion reward as least as  $TR(s)$ . If some of the nodes that are included in the  $Path(s)$  are willing to additionally pass the marketing information to other people who are not included in the  $Path(s)$ , the real diffusion reward will be greater than  $TR(s)$ .

Note that the model could be easily extended to multiple paths starting from node  $s$ . For example, in Figure 1, if we could revise the reward function to use the maximal and submaximal values at the same time for planning the path, the path of  $e_2$  would be extended to multiple paths (starting from  $k_1$  and  $u_4$ , respectively) as shown in Figure 4. However, the diffusion reward would be greater than that in the single-path planning. Generally, the choice of the number of neighbour nodes to forward is determined by the total cost of the incentive to induce message forwarding, which increases as the number of endorsers becomes larger.



**Figure 4.** Example of multiple path planning.

#### 4. Experiment Design

In this section, we apply the proposed mechanism to the micro-blogging system to examine its effectiveness. Micro-blogging services are 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, Plurk is very popular in Asia and the United States [6]. It allows users to send and respond to messages in short sentences (limitation of 140 characters). In addition, it also attracts users to communicate with each other and share external information through embedded URLs. Because Plurk is popular and predominantly used for communicating and sharing, it is an excellent platform for conducting information diffusion while marketers carry out social media marketing.

When studying the issues related to social networks, the snowball sampling method is a feasible approach to use for constructing experiments [3], and it was used in our experiments. First, we randomly selected seven active Plurk users and invited them to join our experiment. In total, 265 participants were generated through the snowball sampling process. After removing the people who were not interested in our experiment, we were left with a total of 131 active Plurk users as participants, and they were also the candidates for the start point of a diffusion path. The data were gathered from the period 10/11/2014 to 16/11/2014. First, to construct the interaction network for obtaining the transition probability, we collected the last 6 months' micro-blogging messages (including post and response data) from participants' public Plurk interface. Then, to construct the friendship network for obtaining the information influenceability and reachability, we recursively expanded friendship from the participants' friends' list. Finally, there were 4,832 social nodes

included in the friendship network. The information regarding the collected social network data is given in Table 1.

**Table 1.** Data descriptions of the experiment

Statistics of the experiment data	
Number of invited participants	131
Average number of friends per participant	37
Average number of active social nodes per participant	11
Average number of monthly interactions per participant (6 months)	2,147

In the experiment, we totally diffused 40 pieces of marketing information through two different marketing strategies: (1) seeking business opportunities and (2) establishing brand expression. According to the previous studies, coupon promotions could cause an increase in product sales [8] and product reviews from third parties might spread good news/impressions of brands that will increase the effectiveness of firms' advertising [16]. There were in total 20 product deals/coupon advertisements for seeking business opportunities and 20 product evaluation review articles for establishing brand expression. The former marketing information was collected from Yahoo! Shopping, which is one of the largest online shopping sites, and the latter was collected from Epinions, which is one of the most professional and best-known product review platforms allowing users to share their product experiences and opinions. To perform the preference fitness analysis, the keywords that most represent the product are needed. In our experiments, the keywords for marketing information were given by an expert group made of six senior graduate students and four doctoral students at the business colleges. The advertisements were delivered with an online 5-star rating questionnaire for the marketing information receivers to feed back their acceptance and diffusion path tracking (Which friend was the marketing information received from?).

We evaluated our proposed mechanism by comparing with the following benchmark approaches: (1) random advertising without a path planning mechanism (Random), (2) random advertising with a path planning mechanism (Random+Path), (3) influencer advertising without a path planning

mechanism (Influencer) and (4) influencer advertising with a path planning mechanism (Influencer+Path). According to Kiss and Bichler [35], out-degree centrality has a better performance in influencer identification, so we used out-degree influencer selection to select the starting point of information diffusion. In addition, the random advertising method randomly selected participants whose  $sb(U) > 0$  as the start point of the information diffusion process. For each advertising method, we selected five participants as starting points for diffusing the marketing information.

## 5. Experiment Results

To evaluate the performance of different advertising methods, we used the advertisement click-through rate (CTR) and the receivers' 5-star acceptance rating feedback on the received marketing message as the evaluation indicators. The former is a popular practical indicator about advertising efficiency; the latter can evaluate the user's impression of the received marketing message.

Intuitively, for seeking business opportunities, it is expected to seek potential customers with high ( $\geq 4$ -star) acceptance of the product advertisement, and for establishing brand expression, it is expected to seek potential customers with not the lowest ( $\geq 2$ -star) acceptance of the product advertisement. We compare the performance using CTR with different star rating conditions.

### 5.1. Seeking Business Opportunities Strategy

Generally, business opportunities exist in potential customers with high acceptance of the product advertisement, which means that he/she has a higher likelihood of buying products. The CTR with acceptance condition formula is defined as

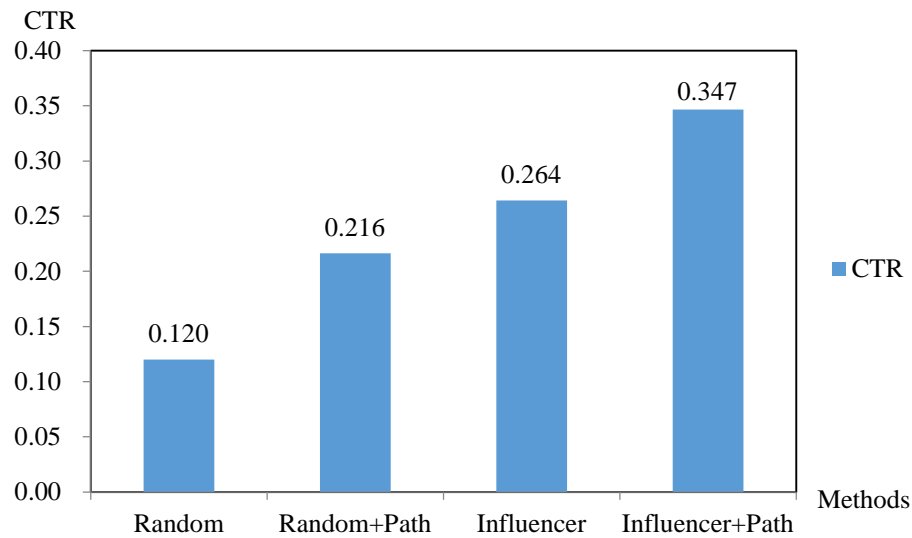
$$CTR = \frac{|\Phi_{\text{click}} \cap \Phi_{4\text{-star}}|}{|\Phi_{\text{ad}}|},$$

where  $|\Phi_{\text{ad}}|$  denotes the total number of delivered advertisements,  $|\Phi_{\text{click}}|$  denotes the total number of clicked/read advertisements and  $|\Phi_{4\text{-star}}|$  denotes the total number of receiver ratings  $\geq 4$ -star acceptance.

In the experiments, the advertisements could be forwarded within four steps from the starting node. Figure 5 shows the CTRs of each step with respect to different benchmark methods. After four



steps forward, the 20 pieces of advertisements in ‘Random’ and ‘Random+Path’ diffused 583 and 776 times in total and got 0.120 and 0.216 CTR, respectively, which means that our path planning mechanism improved the chance of seeking business opportunities by approximately 10%. The advertisements in ‘Influencer’ and ‘Influencer+Path’ diffused 852 and 1,067 times in total and got 0.264 and 0.347 CTR, respectively, which means that our path planning mechanism improved the chance of seeking business opportunities by approximately 8%.



**Figure 5.** CTR in seeking business opportunities.

Furthermore, a 95% significance level two-paired sample t-test was used to evaluate the overall performance of different advertising strategies. The results are shown in Table 2. First, the test results show that the proposed path planning mechanism significantly improves the benchmark advertising methods. In addition, the diffusion effectiveness is also significantly improved if the path planning starts from qualified start points.

**Table 2.** Statistical verification of the CTR under seeking business opportunities strategy

Paired group	Mean	Std deviation	Std error mean	T value	Sig. (2-tailed)
<b>Random+Path vs Random</b>	0.098	0.123	0.027	3.604	0.002
<b>Influencer+Path vs Influencer</b>	0.109	0.202	0.045	2.392	0.027
<b>Influencer+Path vs Random+Path</b>	0.172	0.224	0.050	3.433	0.003

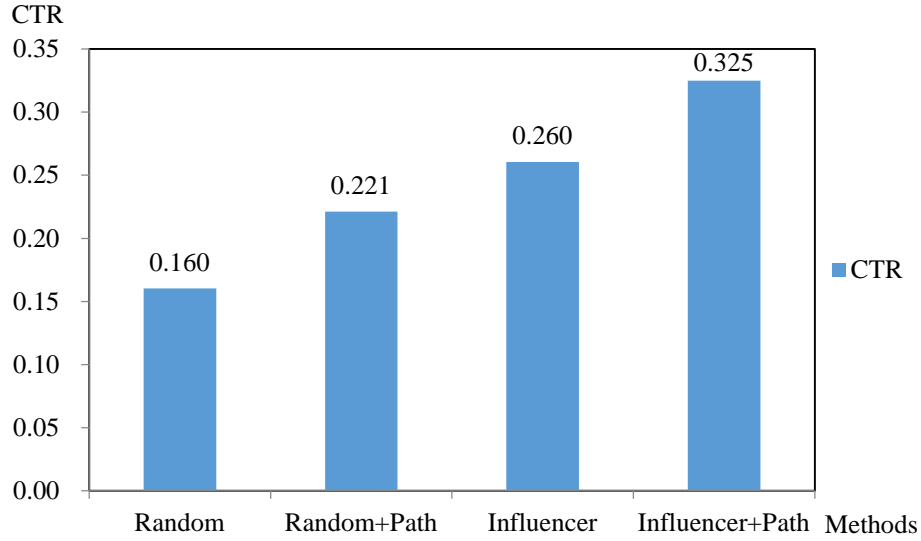
## 5.2. Establishing Brand Expression Strategy

The purpose of this marketing strategy is to enhance (4–5 stars) or reverse (2–3 stars) the brand expression of customers. However, it is very hard to reverse the brand expression of antis (0–1 star). Nevertheless, it might have the opposite effect on marketing strategies. The CTR with acceptance condition formula is defined as

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

where  $|\Phi_{\text{ad}}|$  denotes the total number of delivered advertisements,  $|\Phi_{\text{click}}|$  is the total number of clicked/read advertisements and  $|\Phi_{2\text{-star}}|$  denotes the total number of receiver ratings  $\geq 2$ -star acceptance.

Figure 6 shows the CTR in different benchmark methods. The 20 pieces of advertisements in ‘Random’ and ‘Random+Path’ diffused 985 and 1,243 times in total and got 0.160 and 0.221 CTR, respectively, which means that our path planning mechanism improved the chance of establishing brand expression by approximately 6%. The advertisements in ‘Influencer’ and ‘Influencer+Path’ diffused 1,601 and 1,887 times in total and got 0.252 and 0.321 CTR, respectively, which means that our path planning mechanism improved the chance of establishing brand expression by approximately 7%. Finally, the result of the overall performance of different approaches is further evaluated by two-paired sample t-test and shown in Table 3. At the 95% significance level, all the test results showed that the proposed path planning mechanism significantly improves the other advertising approaches.



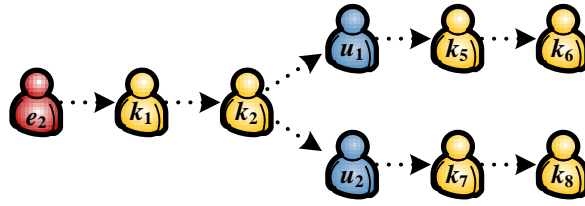
**Figure 6.** CTR in establishing brand expression.

**Table 3.** Statistical verification of the CTR under establishing brand expression strategy

Paired group	Mean	Std deviation	Std error mean	T value	Sig. (2-tailed)
Random+Path vs Random	0.074	0.131	0.029	2.529	0.020
Influencer+Path vs Influencer	0.072	0.124	0.027	2.609	0.017
Influencer+Path vs Random+Path	0.076	0.106	0.023	3.206	0.005

### 5.3. Exposure Ability in Different Strategies

Advertisers are concerned about effective exposure for their advertisements. The proposed APPM would plan a suitable diffusion path for advertisements in different strategies. In one of diffusion, 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 shown in Figure 1, the nodes  $u_1$ ,  $u_2$ ,  $u_3$  and  $u_4$  are the exposure range of the planned diffusion path. Because the path was broken by node  $k_2$  ( $k_2$  delivers the marketing information to nodes  $u_1$  and  $u_2$  rather than the planned node  $k_3$ ) and the system respectively replans the diffusion path for  $u_1$  and  $u_2$ , the planned diffusion paths of the diffusion would be adjusted as shown in Figure 7.



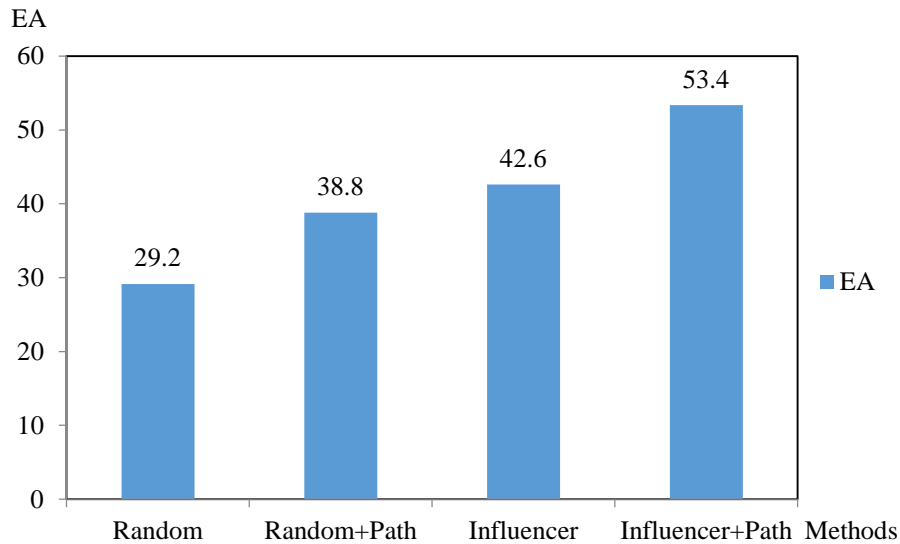
**Figure 7.** Adjusted diffusion path.

However, the replanned diffusion paths still belong to the same marketing information diffusion process. The eventual number of diffusion message receivers is an important indicator in evaluating the performance of the planned diffusion path. The exposure ability (EA) is the average number of receivers per marketing information and is formulated as follows:

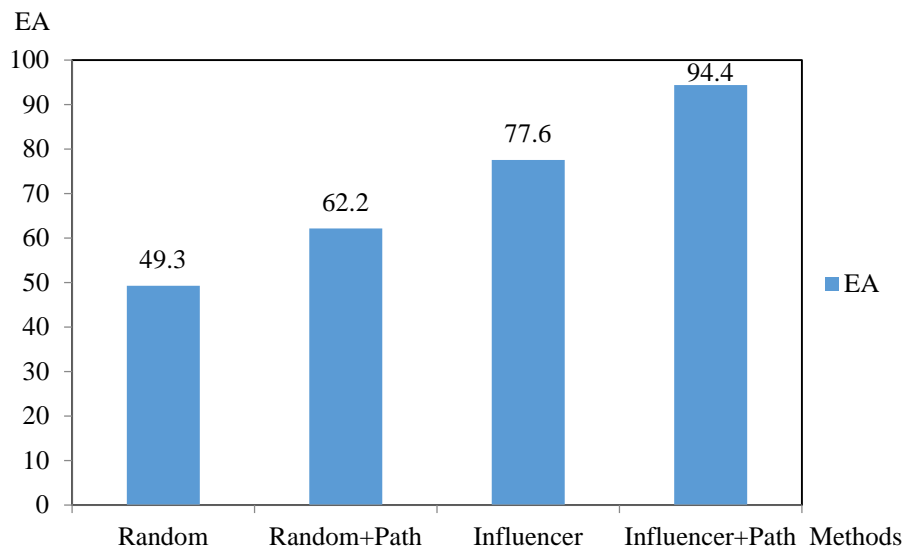
$$EA = \frac{|\Phi_{\text{receivers}}|}{|\Phi_{\text{mi}}|},$$

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

From Figures 8 and 9, we can see that the proposed APPM could enhance the exposure ability of product advertisements if we ignore the acceptance of product advertisements. For the random advertising method, after forwarding for four steps, APPM improves the exposure ability of the random advertising method in the seeking business opportunities strategy and in the establishing brand expression strategy by approximately 33% and 26%, respectively. For the influencer advertising method, APPM improves the exposure ability of the random advertising method in the seeking business opportunities strategy and in the establishing brand expression strategy by approximately 25% and 22%, respectively.



**Figure 8.** Exposure ability in seeking business opportunities strategy.



**Figure 9.** Exposure ability in establishing brand expression strategy.

Here, the paired sample t-test was also performed to further confirm the significant difference of the results of benchmark approaches under different strategies as shown in Tables 4 and 5. At 95% significance level, all the test results show that the advertising strategies with APPM significantly outperformed the advertising strategies without APPM. Therefore, it proves that our proposed strategy is better than the other strategies.

**Table 4.** Statistical verification of the EA under seeking business opportunities strategy

Paired group	Mean	Std	Std error	T value	Sig.
--------------	------	-----	-----------	---------	------

	Deviation	Mean		(2-tailed)
<b>Random+Path vs Random</b>	9.65	14.01	3.13	3.080
<b>Influencer+Path vs Influencer</b>	10.75	18.81	4.21	2.556
<b>Influencer+Path vs Random+Path</b>	14.55	17.36	3.88	3.748

**Table 5.** Statistical verification of the EA under establishing brand expression strategy

Paired group	Mean	Std deviation	Std error mean	T value	Sig. (2-tailed)
<b>Random+Path vs Random</b>	12.90	10.03	2.244	5.748	0.000
<b>Influencer+Path vs Influencer</b>	16.80	19.71	4.406	3.813	0.001
<b>Influencer+Path vs Random+Path</b>	32.20	20.06	4.487	7.176	0.000

#### 5.4. Sharing Behaviour Evaluation

This section further evaluates the sharing behaviours in different advertisement diffusion processes. As mentioned before, egoism and altruism are the two significant factors of willingness-to-share behaviour. There are four delivering situations discussed as shown in Table 6:

- (1) Indicating that the forwarder expects to get positive recognition from receivers. It is most beneficial to both business opportunities seeking and brand expression establishing strategies.
- (2) Indicating that the forwarder expects to influence the impression of receivers of a specific product/brand. It may be helpful to the brand expression establishing strategy.
- (3) Indicating that the forwarder expects to inform the receivers about some promotion information about products. It is most beneficial to the business opportunities seeking strategy.
- (4) Although this could also indicate that the forwarders expect to get negative recognition from receivers, it has no benefits for business. Furthermore, it is possibly just blind delivery behaviour. It is defined as ineffective propagation.

**Table 6.** Statistical verification of the EA under establishing brand expression strategy

		Receiver	
		High preference	Low preference
Forwarder	High preference	(1) Egoism	(2) Altruism for establishing brand expression
	Low preference	(3) Altruism for seeking business opportunities	(4) Ineffective propagation

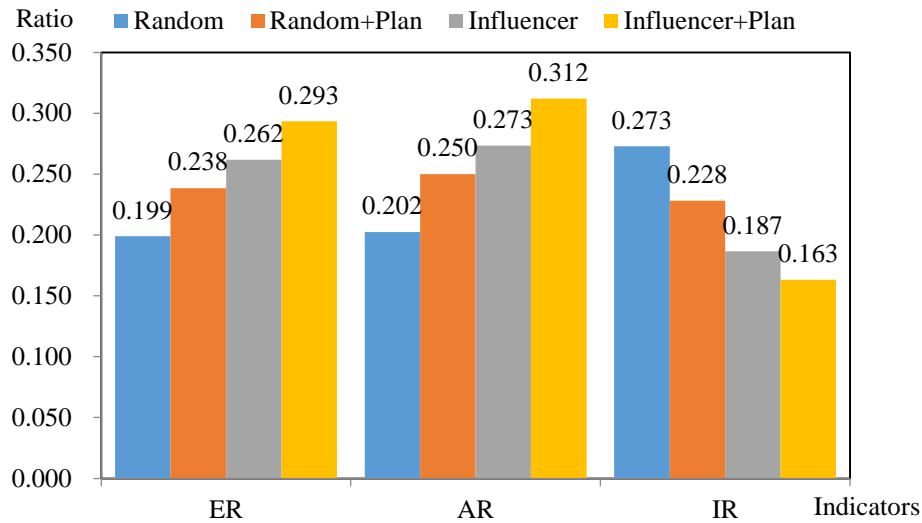
Then, we define the egoism ratio (ER), altruism ratio (AR) and ineffective deliver ratio (IR) for each advertisement diffusion process as shown in Table 7 for evaluating whether the APPM could take advantage of egoism and altruism sharing motivations. In the formulations, we simply define the high preference value ( $PF^H$ ) as  $PF \geq 0.5$  and the low preference value ( $PF^L$ ) as  $PF < 0.5$ .

**Table 7.** Data descriptions of the experiment

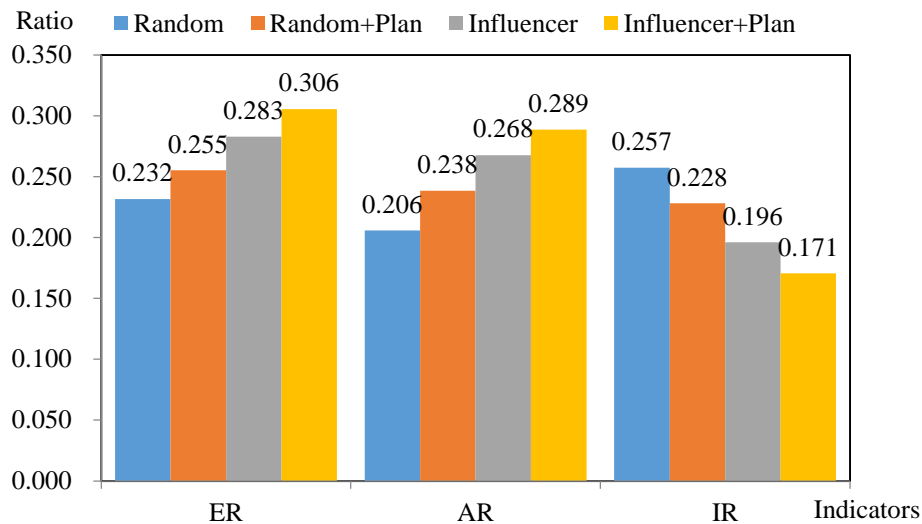
	Seeking business opportunities strategy	Establishing brand expression strategy
<b>Egoism ratio</b>	$ER = \frac{ \Phi_{\text{forwarder} \cap PF^H} \cap \Phi_{\text{receiver} \cap PF^H} }{ \Phi_{\text{forwards}} }$	
<b>Altruism ratio</b>	$AR = \frac{ \Phi_{\text{forwarder} \cap PF^L} \cap \Phi_{\text{receiver} \cap PF^H} }{ \Phi_{\text{forwards}} }$	$AR = \frac{ \Phi_{\text{forwarder} \cap PF^H} \cap \Phi_{\text{receiver} \cap PF^L} }{ \Phi_{\text{forwards}} }$
<b>Ineffective delivery ratio</b>	$IR = \frac{ \Phi_{\text{forwarder} \cap PF^L} \cap \Phi_{\text{receiver} \cap PF^L} }{ \Phi_{\text{forwards}} }$	

where  $|\Phi_{\text{forwards}}|$  denotes the total times of forwards of the advertisement,  $|\Phi_{\text{forwarder} \cap PF^H}|$  and  $|\Phi_{\text{forwarder} \cap PF^L}|$  are the total number of forwarders who have high and low preference fitness, respectively, and  $|\Phi_{\text{receiver} \cap PF^H}|$  and  $|\Phi_{\text{receiver} \cap PF^L}|$  denote the total number of high and low preference

fitness receivers, respectively, who receive the advertisement from the forwarders.



**Figure 10.** Sharing behaviour evaluations in seeking business opportunities strategy.



**Figure 11.** Sharing behaviour evaluations in establishing brand expression strategy.

From Figures 10 and 11, it can easily be seen that the proposed APPM could take advantage of egoism and altruism sharing motivations and decrease the ineffective delivery ratio in both strategies. In addition, we found that all the ARs are higher than ERs in the seeking business opportunities strategy. This indicates that the altruism-motivated users [with higher value of  $sb(\cdot)$ ] are helpful for business opportunities seeking. Because of this, if the altruism-motivated users do not have a preference regarding the information, they are still willing to share the information with friends who might like it. In the brand expression establishing strategy, all the ERs are higher than ARs, which



means that the egoism-motivated users [with higher value of  $PF(\cdot)$ ] are more beneficial for establishing brand expression. Because the egoism-motivated users expect to obtain responses and reputations, they are willing to share the information that they know and are interested in.

## 6. Discussion

In this paper, we proposed an APPM, which is based on probability and optimization models. Our mechanism treats the diffusion problem as a sequential optimization problem. We incorporate preference fitness analyzing, transition flow inferring, and customer value evaluating and diffusion path planning techniques to plan the optimal diffusion path for influential social nodes. To identify the transition probability of the possible transition states, we first constructed an interaction network based on the daily social interactions within a social network. Then, the LSI-based methodology is applied to identify the preference fitness of users from their daily micro-blogging messages. The concept of the Markov chain is used to derive the transition probabilities between the active social nodes. Social network analysis based on the constructed interaction network is adopted to obtain the information influenceability and the reachability of social nodes. Finally, a simple probability model consolidating the other submodules is used to calculate the expected value of path planning.

There are several limitations to this research. First, because of the privacy issue, it was difficult to extract online personal data (e.g. social information). Therefore, we invited participants to join in the experiments. If more users are recruited and engaged, the accuracy of the proposed mechanisms will be better. Second, the ratio-based determination has a possibility of data bias regarding the frequency of use in the period of data collection. The directions of trustworthiness and social influence between users could be taken into consideration. When determining the possible transition states and the transition probabilities, the concept of trust and the tie strength analysis between social nodes might reflect reality more. Third, in the current paper, online postings in social media were used as social interactions for analyzing the strength of interpersonal relationships. In social media, there are many ways (e.g. messaging, applications, photo uploads, and chat) for users to interact with others. The analysis of relationship strength would be more comprehensive if other interaction methods were considered. Lastly, the experiments were conducted on a single micro-blogging

platform. The effect of the platform on the users' diffusion behaviours could be further examined and compared if multiple types of social network platforms were considered.

## 7. Conclusion

In this paper, we proposed an APPM, which is based on probability and optimization models. Our mechanism treats the diffusion problem as a sequential optimization problem. We incorporate preference fitness analyzing, transition flow inferring, and customer value evaluating and diffusion path planning techniques to plan the optimal diffusion path for influential social nodes. To identify the transition probability of the possible transition states, we first constructed an interaction network based on the daily social interactions within a social network. Then, the LSI-based methodology is applied to identify the preference fitness of users from their daily micro-blogging messages. The concept of the Markov chain is used to derive the transition probabilities between the active social nodes. Social network analysis based on the constructed interaction network is adopted to obtain the information influenceability and the reachability of social nodes. Finally, a simple probability model consolidating the other submodules is used to calculate the expected value of path planning. Our experimental results show that the proposed mechanism outperforms other benchmark approaches and can significantly improve the effectiveness of advertising message diffusion.

The contributions and managerial implications of this research are summarized as follows. First, from the perspective of system innovation, although social media marketing has become increasingly popular, little research has proposed a diffusion planning mechanism to help the influencers boost their advertising effectiveness for propagating information. We are one of the pioneers in treating the information diffusion problem as a sequential path planning optimization problem rather than a simple influential node recommendation issue. Second, from the perspective of methodology, we consider not only the individual preference and social influence (influenceability and reachability) but also behavioural factors (interaction transition probability and willingness to share) in the evaluation of the reward function to identify the path that could gain the maximum diffusion reward. Third, from the perspective of performance, the evaluation results confirm 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 in random influencer selection for selecting diffusion start points, the proposed path planning mechanism could support and improve the diffusion effectiveness, and the mechanism would be able to achieve a greater performance if it was combined with other influencer discovery mechanisms. Lastly, from the perspective of practice, the mechanism can help marketers to conservatively evaluate the possible information diffusion effectiveness under different marketing strategies and support the evaluated influencers propagating information to specific individuals to continue the diffusion process. Furthermore, the proposed mechanism could take advantage of both egoism and altruism sharing motivations and decrease the ineffective delivery ratio under different marketing strategies.

There are some aspects that can be further improved. First, the optimal path planning formulation might be subject to some conditions, for example, both the influenceability and reachability of social nodes should be greater than a threshold according to the marketers' needs. The determination of the threshold to improve the planning effectiveness could be further investigated. Second, different social factors could be taken into consideration when formulating the diffusion reward function. For example, if a social node is located in a structural hole, the marketer might gain relatively great diffusion reward from him/her. Third, the impact of preference fitness and sharing behaviour indicators on the path planning module could be further examined. The effectiveness of the different marketing strategies should be improved if these two indicators can be combined appropriately. Finally, social network-based mechanisms generally investigate novel online services from many perspectives (e.g. social perspective, structural and behavioural factors, personal and group characteristics, and public and private information). Different multiple criteria decision methods for balancing varied indicators could be applied to improve the effectiveness of the diffusion mechanism.

## References

- [1] M.A. Abubakre, M.N. Ravishankar, Crispin R. Coombs, The role of formal controls in facilitating information system diffusion, *Information & Management* 52 (2015) 599–609.
- [2] E. Adar and L.A. Adamic, Tracking information epidemics in blogspace, in: *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Web Intelligence, 2005*, IEEE Computer Society, Washington, DC, USA.
- [3] Y.Y. Ahn, S. Han, H. Kwak, S. Moon, H., Jeong, Analysis of topological characteristics of huge online social networking services, in: C. Williamson, M.E. Zurko, P. Patel-Schneider, P. Shenoy (Eds.), *Proceedings of the 16th International Conference on World Wide Web*. New York, ACM Press, 2007, pp. 835–844.
- [4] PwC and IAB, *IAB Internet Advertising Revenue Report – An Industry Survey (2011 Full Year Results)*, PricewaterhouseCoopers and Interactive Advertising Bureau, April 2012.
- [5] K. Bagwell, *The Economic Analysis of Advertising*, *Handbook of Industrial Organization* Volume 3, Edward Elgar Press, Cheltenham, UK, 2007.
- [6] Bangalore, Analysis on the general profile of users on Plurk.com, InRev systems. <http://www.slideshare.net/bexdeep/plurk-analysis-4136802> (accessed 20.2. 12).
- [7] M.S. Balaji, K.W. Khong, A.Y.L Chong, Determinants of negative word-of-mouth communication using social networking sites, *Information & Management* 53(4) 2016, 528-540.
- [8] K. Bawa, R.W. Shoemaker, Analyzing incremental sales from a direct mail coupon promotion, *Journal of Marketing* 53 (3) (1989), 66–78.
- [9] M. Berry, S. Dumais, G. O'Brien, Using linear algebra for intelligent information retrieval, *SIAM Review* 37 (4) (1995), 573–595.
- [10] S.P. Borgatti, Identifying sets of key players in a network, *Compute Math Organ Theory* 12 (1) (2006), 21–34.
- [11] S.P. Borgatti, The key player problem, in: *Proceeding of the Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, ed. by R. Breiger, K. Carley, P. Pattison. Washington, DC: Committee on Human Factors, National Research Council, 2003.

- [12] D. Brown, N. Hayes, Chapter 20: Making influencer marketing work for your company, *Influencer Marketing*, Butterworth-Heinemann, Oxford, 2008, 212–223.
- [13] J. Cadzow, SVD representation of unitarily invariant matrices, *IEEE Transactions on Acoustics, Speech and Signal Processing* 32 (3) (1984), 512–516.
- [14] Y.H. Chen, T.P. Lin, D.C. Yen, How to facilitate inter-organizational knowledge sharing: The impact of trust, *Information & Management* 51(5) 2014, 568-578.
- [15] J. Chen, X.L. Shen, Consumers' decisions in social commerce context: An empirical investigation, *Decision Support Systems*, 79 (2015) 55-64.
- [16] Y.B. Chen, J.H. Xie, Third-party product review and firm marketing strategy, *Marketing Science* 24 (2) (2005), 218–240.
- [17] Y. Chen, F.S. Tsai, K.L. Chan, Machine learning techniques for business blog search and mining, *Expert Systems with Applications* 35 (3) (2008), 581–590.
- [18] J.A. Chevalier, D. Mayzlin, The effect of word of mouth on sales: online book reviews, *Journal of Marketing Research* 43 (3) (2006), 345–354.
- [19] Y. Cho, J. Hwang, D. Lee, Identification of effective opinion leaders in the diffusion of technological innovation: a social network approach, *Technological Forecasting and Social Change* 79 (1) (2012), 97–106.
- [20] S. Deerwester, S. Dumais, G. Furnas, T. Landauer, R. Harshman, Indexing by latent semantic analysis, *Journal of the American Society for Information Science* 41 (6) (1990), 391–407.
- [21] R. Dunbar, *How Many Friends Does One Person Need? Dunbar's Number and Other Evolutionary Quirks*, Harvard University Press Cambridge, Massachusetts 2010.
- [22] M. Eisend, Source credibility dimensions in marketing communication: a generalized solution, *Journal of Empirical Generalizations in Marketing* 10 (2006), 1–33.
- [23] N.E. Friedkin, E.C. Johnsen, *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*, Cambridge, Cambridge University Press, 2011.
- [24] D. Gleich, L. Zhukov, SVD-based term suggestion and ranking system, in: *Proceedings of the 4th IEEE International Conference on Data Mining*, 2004, 391–394.

- [25] D., Godes, D. Mayzlin, Using online conversations to study word-of-mouth communication, *Marketing Science* 23 (4) (2004), 545–560.
- [26] R.E. Hanneman, Introduction to Social Network Methods, Online Textbook Supporting Sociology, 2000, Riverside, CA, University of California.
- [27] D.L. Hoffman, M. Fodor, Can you measure the ROI of your social media marketing? *MIT Sloan Management Review* 52 (1) (2010), 41–49.
- [28] C.L. Hsu, C.C. Lin, Acceptance of blog usage: the roles of technology acceptance, social influence and knowledge sharing motivation, *Information & Management* 45 (1) (2008), 65–74.
- [29] B.A. Huberman, D.M. Romero, F. Wu, Social networks that matter twitter under the microscope, *First Monday*, 14 (1–5) (2009).
- [30] G. Iyer, D. Soberman, J.M. Villas-Boas, The targeting of advertising, *Marketing Science* 24 (3) (2005), 461–476.
- [31] A. Jackson, J. Yates, W. Orlikowski, Corporate blogging: building community through persistent digital talk, in: *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, 2007, Washington, DC, USA.
- [32] P. Kazienko, M. Adamski, AdROSA: adaptive personalization of web advertising, *Information Sciences* 177 (11) (2007), 2269–2295.
- [33] A.J. Kim, E. Ko, Do social media marketing activities enhance customer equity? An empirical study of luxury fashion brand, *Journal of Business Research* 65 (10) (2012) 1480–1486.
- [34] D. Kim, T. Ammeter, Predicting personal information system adoption using an integrated diffusion model, *Information & Management* 51 (2014) 451–464.
- [35] C. Kiss, M. Bichler, Identification of influencers: measuring influence in customer networks, *Decision Support Systems* 46 (1) (2008), 233–253.
- [36] A. Kontostathis, W.M. Pottenger, A framework for understanding Latent Semantic Indexing (LSI) performance, *Information Processing & Management* 42 (1) (2006), 56–73.
- [37] L.I. Labrecque, E. Markos, G.R. Milne, Online personal branding: processes, challenges, and implications, *Journal of Interactive Marketing* 25 (1) (2011), 37–50.

- [38] M. Laroche, M.R. Habibi, M.O. Richard, R. Sankaranarayanan, The effects of social media based brand communities on brand community markers, value creation practices, brand trust and brand loyalty, *Computers in Human Behavior* 28 (5) (2012), 1755-1767.
- [39] Y.M. Li, Y.L. Shiu, A diffusion mechanism for social advertising over microblogs, *Decision Support Systems* 54 (1) (2012), 9-22.
- [40] Y.M. Li, C.Y. Lai, C.W. Chen, Discovering influencers for marketing in the blogosphere, *Information Sciences* 181 (23) (2011), 5143–5157.
- [41] L.F. Lin, Y.M. Li, W.H. Wu, A social endorsing mechanism for target advertisement diffusion, *Information & Management*, 42 (2015) 982–997.
- [42] Z. Liu , Q. Min, Q. Zhai, Russell Smyth, Self-disclosure in Chinese micro-blogging: a social exchange theory perspective, *Information & Management* 53 (2016) 53–63.
- [43] A. Mislove, M. Marcon, K.P. Gummadi, P. Druschel, B. Bhattacharjee, Measurement and analysis of online social networks, in: *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, 2007, ACM, New York, NY, USA.
- [44] M. Newman, The structure and function of complex networks, *SIAM Review* 45 (2) (2003), 167–256.
- [45] Nielsen, State of the media: the social media report q3 2011. <http://blog.nielsen.com/nielsenwire/social/> (accessed 30.05. 12).
- [46] Y.G. Pan , Y. Xu , X. Wang, C.H. Zhang , H. Ling , J. Lin, Integrating social networking support for dyadic knowledge exchange: a study in a virtual community of practice, *Information & Management* 52 (2015) 61–70.
- [47] D.H. Park, S. Kim, The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews, *Electronic Commerce Research and Applications* 7 (4) (2008), 399–410.
- [48] M. Risius, R. Beck, Effectiveness of corporate social media activities in increasing relational outcomes, *Information & Management* 52(7) 2015, 824-839.
- [49] B. Ruhnau, Eigenvector-centrality: a node-centrality?, *Social Networks* 22 (4) (2000), 357–365.

- [50] G. Salton, A. Wong, C.S. Yang, A vector space model for automatic indexing, *Communications of the ACM* 18 (11) (1975), 613–620.
- [51] M. Stelzner, Social media marketing industry report. <http://www.socialmediaexaminer.com/social-media-marketing-industry-report-2012/> (accessed 30.05.12).
- [52] I. Roelens, P. Baecke, D.F. Benoit, Identifying influencers in a social network: The value of real referral data, *Decision Support Systems* (2016), <http://dx.doi.org/10.1016/j.dss.2016.07.005>.
- [53] K. Wang, E.T.G. Wang, C.K. Farn, Influence of Web advertising strategies, consumer goal-directedness, and consumer involvement on Web advertising effectiveness, *International Journal of Electronic Commerce* 13 (4) (2009), 67–96.
- [54] C. Wen, B.C.Y. Tan, K.T.T. Chang, Advertising effectiveness on social network sites: an investigation of tie strength, endorser expertise and product type on consumer purchase intention, in: *Proceedings of the International Conference on Information Systems, 2009*, Phoenix, Arizona, United States.
- [55] DEI Worldwide and online testing exchange, engaging consumers online: the impact of social media on purchasing behaviour, 2008. [www.deiworldwide.com/files/DEIStudy-Engaging%20ConsumersOnline-Summary.pdf](http://www.deiworldwide.com/files/DEIStudy-Engaging%20ConsumersOnline-Summary.pdf) (accessed 07.06.12).
- [56] Z. Yan, T. Wang, Y. Chen, H. Zhang, Knowledge sharing in online health communities: A social exchange theory perspective, *Information & Management* 53(5) 2016, 643-653.
- [57] J. Yang, J. Leskovec, Modeling information diffusion in implicit networks, in: *Proceedings of the 2010 IEEE International Conference on Data Mining, 2010*, Sydney, NSW.
- [58] W.S. Yang, J.B. Dia, H.C. Cheng, H.T. Lin. Mining social networks for targeted advertising, in: *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, IEEE Computer Society, 2006, Washington, DC, USA.
- [59] Z. Yu, C. Wang, J. Bu, X. Wang, Y. Wu, C. Chen, Friend recommendation with content spread enhancement in social networks, *Information Sciences*, 309 (2015), 102–118.
- [60] K.D. Williams, K.B. Williams, Impact of source strength on two compliance techniques, *Basic*



and *Applied Social Psychology* 10 (2) (1989), 149–159.

- [61] H. Zhang, Y. Lu, S. Gupta, L. Zhao, What motivates customers to participate in social commerce? The impact of technological environments and virtual customer experiences, *Information & Management* 51(8) 2014, 1017-1030.

## Biographies

**Yung-Ming Li** is a Professor at the Institute of Information Management, National Chiao Tung University in Taiwan. He received his Ph.D. in Information Systems from the University of Washington. His research interests include network science, Internet economics and business intelligence. His research has appeared in *IEEE/ACM Transactions on Networking*, *INFORMS Journal on Computing*, *Decision Sciences*, *European Journal of Operational Research*, *Decision Support Systems*, *International Journal of Electronic Commerce*, *Information & Management*, *ICIS*, and *WITS*, among others.

**Cheng-Yang Lai** received his Ph.D. from the Institute of Information Management, National Chiao Tung University in Taiwan. His research interests include electronic commerce and business intelligence. His research has appeared in *Decision Support Systems*, *Electronic Commerce Research and Applications*, and *Information Sciences*.

**Lien-Fa Lin** is an Associate Professor at the Department of Information Communication, Kao Yuan University in Taiwan. He received his Ph.D. in Information Management from National Chiao Tung University. His research interests include electronic commerce, mobile computing, and network economics. His research has appeared in *Decision Support Systems*, *International Journal of Electronic Commerce*, and *Information and Management*.