INDIAN INSTITUTE OF MANAGEMENT CALCUTTA

WORKING PAPER SERIES

WPS No. 733/ October 2013

A graph-based scheme for Brand Promotion in Social Media Platforms using Influencer nodes

by

Divya Sharma
Doctoral student, IIM Calcutta D. H. Road, Joka, Kolkata 700104, India

Debashis Saha
Professor, IIM Calcutta, Diamond Harbour Road, Joka P.O., Kolkata 700 104 India

&

Parthasarathi Dasgupta
Professor, IIM Calcutta, Joka, Diamond Harbour Road, Kolkata 700104, India

A graph-based scheme for Brand Promotion in Social Media Platforms using Influencer nodes

Divya Sharma, Debashis Saha, Parthasarathi Dasgupta

Indian Institute of Management Calcutta

Social media is emerging as the mode of choice among internet users to communicate, collaborate and share content. At the same time, social media is also unfolding as an important platform to allow marketers to reach out to their customers and improve the customer experience by establishing closer ties. In addition, marketers are increasingly using social media to identify potential customers and then influence these leads to purchase their products, subscribe to their services and connect with their brand.

Most social media websites allow their users to create a public or semi-public profile. The profile may contain information about the user such as, demographics (like age, gender, location, educational qualifications, occupation), interests, subscriptions, etc. It is the availability of this type of personal information about its users which have made social media websites a remarkable opportunity to tap into a large pool of potential customers.

Social media websites are now allowing marketers to not only advertise their brand and products, but, also connect with their existing and future customers. Marketers are using multiple social media to connect with their customers. Marketers are creating brand pages on Facebook, establishing a dedicated channel on Youtube, tweeting to their followers on Twitter, maintaining their own websites and blogs, and using every means possible to connect with their customers.

Social media websites also provide an opportunity for their users to subscribe to the pages, channels and feeds of their favorite brands. Through such subscriptions, the user can keep track of the latest activity with respect to their subscribed brand. The marketers can post ads directed to their subscribers and can make exclusive offers and announcements to these subscribers.

Advertising on social media websites is also maturing and there have emerged multiple ways of publishing advertisements. For example, Facebook allows advertisers to create relevant ads and sponsored stories (facebook.com). The relevant ads are shown to a user on the right side of the screen or in the newsfeed on the basis of demographic information or recent activity by the user or her friends. The sponsored stories appear in the news feed of a user based on certain activities of her friends on a particular brand pages. When a user subscribes to a brand page or performs an activity on the brand page, this information may be posted to the friends of the subscriber in the form of sponsored stories if the advertiser feels that it would be relevant to their business.

An important facet of using social media for advertising is the scope for utilizing word-of-mouth and endorsements from existing customers and subscribers. In this way, advertisers can identify users who can influence other users to buy, try or subscribe to their products or services. In fact, marketers are trying to harness word-of-mouth by analyzing the connections between members of a network (Katona, 2011). Social networks can, therefore, help advertisers and marketers to not only target potential customers who might be more interested in their products, but also use endorsements and actions of existing customers on their brand pages to influence other users in the friend circles of these customers.

However, there is no limit to how much can be posted on a user's newsfeed, and at the same time, excessive advertising on the newsfeed can become a nuisance for users and spoil their experience. Advertisement clutter can lead to reduced attention by consumers since, firstly consumers may get irritated with the advertisements and subsequently start avoiding the ads, and secondly, limited memory will not allow the consumers to remember all the ads shown to them(Ha, 2008). Since the newsfeed also acts as an advertisement medium, it becomes a challenge for social media websites to ensure an optimal newsfeed for the users that maximizes gains for the advertisers. The focus of this work is to address this challenge, by optimizing the appearance of ads by utilizing word-of-mouth effect. In addition, the attempt is at ensuring that the ads appear in the newsfeed of only those users who might be interested in the brand. As a result of such optimized newsfeed, the intent is to increase the reach of the particular brand page and, at the same time, ascertaining that the users who subscribe to the brand page as a result of social media driven word-of-mouth possess the characteristics desirable in the target audience of the brand.

The Model

For the purpose of this work a social network is visualized as a graph of two types of nodes, namely, user nodes and brand nodes. Every user of the social network is represented as a user node in the graph and advertisers and marketers on the social network are represented through brand nodes. Each user node is linked to a public or semi-public profile and each brand node maintains a public brand page. While the profile of a user contains information about the demographics, interests, subscriptions and activities of the user, the brand page contains information about the particular brand or product of the marketer in the form of multimedia content and a trail of activities on the page by the marketer or subscribers of the page. Each brand page also has an option allowing a visitor to the page to subscribe to it if the page is of interest to her.

For notational convenience, user nodes are depicted as circles and brand nodes are depicted as squares in the network. An edge between two user nodes means that these two users are friends. On the other hand a link between a brand page and a user node means that the user is subscribed to the brand page.The subscribers of the brand page are considered to be at the primary level with respect to the brand page, while the friends of these subscribers who are not subscribed to the brand page are considered to be at the secondary level with respect to the brand page. For the purpose of this work, edges between brand pages are not considered. We consider a sub-graph of a social network with one brand page *b* which has n subscribers denoted by user nodes $s_i$where i = 1, 2, …, n. The friend circle of a user node $s_i$with m friends is depicted by user nodes $s_{ij}$where j = 1, 2, …, m.

Identification of Influencers

As discussed above, social networks provide advertisers and marketers with the opportunity to use word-of-mouth of their existing customer base to influence potential customers. The identification of potential customers is an important task for advertisers and marketers. Social networks can leverage the information about the profiles of its user base to provide a targeting service to advertisers and marketers where the social network searches for the potential customers on behalf of the advertisers and marketers. In this work, the aim is to identify potential customers who are a part of the friend circle of an existing subscriber of a brand page. The social network can identify a desired number, M, of such potential customers and can then display advertisements of the brand page in the newsfeed of these potential customers along with details about recent activity by subscribers who are in the friend circle of these potential customers. In this way, the existing subscribers of the brand page act as influencers for the potential customers.

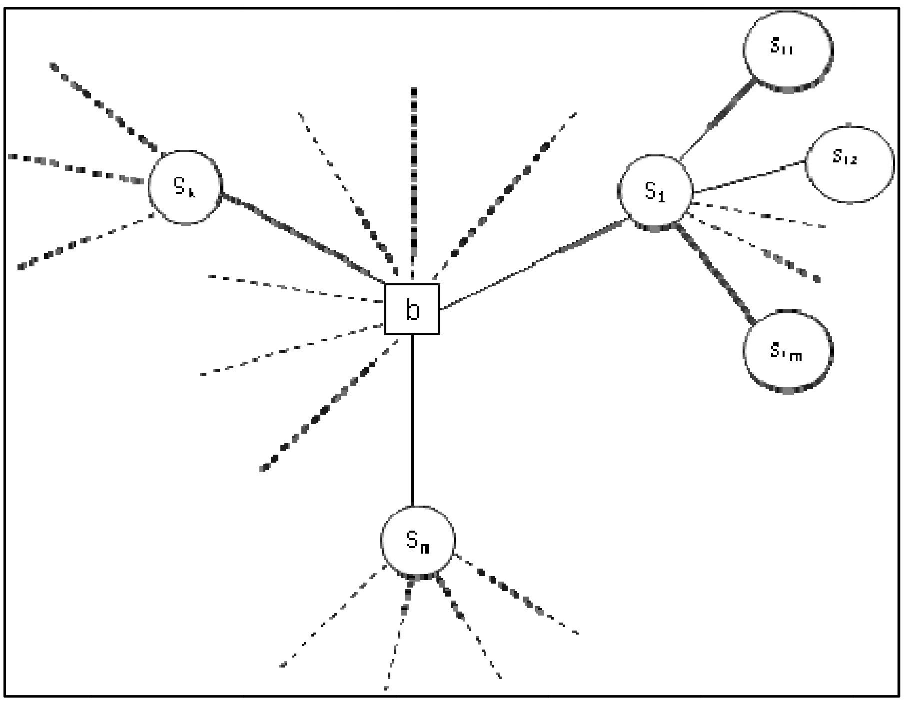

Figure 1 Sub-network of brand page b

Three scenarios for identification of such potential customers may be considered. Figure 1 is used to describe each of these scenarios.

Scenario 1: Using all subscribers as influencers

In this scenario, the social network considers all the subscribers, S_i, of a particular brand page as influencers, F_i. Next, for each influencer the association value between the influencer node, F_i, and each of her friends, S_{ij}, is calculated using the algorithm specified in (Sharma, Alam, Dasgupta, & Saha, 2013). Once the association value is calculated for all friends of all influencers, these values are ranked in descending order of association value. Thereafter, nodes with the top M association values are identified by the social network website and these nodes then act as the potential customers whose newsfeed receives posts containing ads of the brand page along with details about the recent activities of the subscribers who are already part of the friend circle of these potential customers.

The idea behind this scenario is that it makes sense for the advertisers and marketers to target potential customers who are closely linked with their existing customer base, have similar profiles as their existing customer base and interact with their existing customer base so as to allow word-of-mouth marketing. The algorithm in (Sharma, Alam, Dasgupta, & Saha, 2013) captures all these aspects and is, therefore, appropriate for use in this scenario. However, in cases where a brand page has a large subscriber base, using this algorithm will become extremely complex, since, association values will have to be computed for all the friends of all the subscribers to choose the top M potential customers. Also, it is possible that the potential customers identified are friends with existing subscribers who are not as closely associated with the brand page as other subscribers whose friends had comparatively lower association values. In this case, it is likely that the influence of the subscriber will be little.As a result, there will not be much activity between the subscriber and the brand page which can be leveraged for marketing. Also, these potential customers might not show any interest in the ads of the brand page and may also feel that their user experience is getting spoilt by such posts in their newsfeed.

Scenario 2: Using a subset of subscribers as influencers

The shortcoming of the scenario 1 can be overcome by calculating the association value between a brand page and its set of subscribers. The following section discusses the procedure for calculating this association value.

The Metrics

The link between a brand page and a subscriber is weighted by an association value. The association value is, in turn, calculated using three metrics, namely, proximity, similarity and interaction. The definition of each of these metrics is discussed below.

Proximity: Proximity $p_{bi}$is a measure of how closely a particular user *i*is linked to a brand page*b*. In other words, this metric tries to capture the extent to which a user likes a brand page. The closeness that a user feels for a particular brand page is represented as ρ_{bi}.

$$p_{bi} = \rho_{bi} \text{ , where } 0 \le \rho_{bi} \le 1 \text{ and } 1 \le i \le m$$

Similarity: Most social networks allow marketers and advertisers of a brand page to select a set of attributes which are desirable in their set of customers. The set of desirable attributes for brand page *b* is defined as

$$A_b = \{a_{b1}, a_{b2}, \dots, a_{bu}\} \text{ where } a_{bv} \text{ represents the } v^{th} \text{ attribute for brand page } b$$

The social network then uses this set of attributes to find users to whom the ads of a particular brand page must be shown. The advertiser may be allowed to assign weights to the attributes in the set in the order of preference and importance. The weight *w is* assigned to attribute v follows

$$0 < w_v < 1$$

Each user also maintains a profile on the social network which contains information regarding demographics, interests, subscriptions, etc. We define the user profile of a user *i* as

$$P_i = \{d_{i1}, d_{i2}, \dots\}\ where\ d_{ik}\ represents\ the\ k^{th}\ information\ about\ the\ user\ i$$

Similarity is a measure of the extent to which a particular user's profile information matches with the set of desirable attributes defined for a brand page. Similarity $S_{bi}$is mathematically formulated below.

$$S_{bi} = \frac{\sum_1^u w_v e_v}{\sum_1^u w_v}$$

$$where\ e_v = 1, if\ a_v \in P_i$$

$$e_v = 0, otherwise$$

Interaction: Social network websites usually allow subscribers to perform actions on brand page they are subscribed to. Some common forms of interactions that subscribers may have include liking posts of the brand page, sharing posts of the brand page in their own newsfeed and commenting on the posts of the brand page. An interaction between subscriber *i* and brand page *b* of type t is defined in terms of its recency r_{bi} and frequency f_{bi} as

$$\sigma_{bi}^t = \alpha r_{bi}^t + (1 - \alpha) f_{bi}^t$$

$$where\ 0 \le \alpha \le 1$$

α is a measure of the importance of recency in comparison to frequency. An advertiser or marketer may assign more or less weight to recency in relation to frequency depending on the significance of either of them to the brand page.

The recency of an interaction is a measure of how fresh the interaction is. If the association value is being computed at time τ_o and τ_{il}^tcorresponds to the time of the latest interaction of type t between the subscriber *i*and the brand page b then recency is defined as

$$r_{bi}^t = 1 - \frac{\tau_o - \tau_{il}^t}{\omega_l^t}$$

$$where\ \omega_l^t = \max\{\tau_o - \tau_{il}^t \,|\, \forall i, 1 \le i \le n\}$$

The frequency of an interaction is a measure of how often a particular subscriber *i* interacts with the brand page *b*. If V_{il}^t is the volume of interactions of type t which a subscriber has had with a brand page in the time window ω_l^t then frequency is defined as

$$f_{bi}^t = 1 - \frac{1}{V_{il}}$$

The overall interaction between a user *i* and brand page *b* is then defined as a composite of all interactions σ_{bi}^{t} as

$$I_{bi} = \sum_{t} \beta_t \sigma_{bi}^{t}$$

$$where \sum_{t} \beta_t = 1 \; and \; 0 \le \beta_t \le 1$$

β_t represents the weight assigned to interactions of type t. This provision allows the advertiser or marketer to assign different weights to interactions of different types depending on the significance of the different types of interactions to the brand page.

Association:The association value between a subscriber *i* and a brand page *b* can then be defined as

$$A_{bi} = \mu_1 p_{bi} + \mu_2 S_{bi} + \mu_3 I_{bi}$$

$$where \; \mu_1 + \mu_2 + \mu_3 = 1 \; and \; 0 \le \mu_1, \mu_2, \mu_3 \le 1$$

μ_1, μ_2, μ_3 represent the relative significance of proximity, similarity and interaction in calculating the association value.

Scenario 2a: Using a subset of subscribers with association values above a threshold

In this scenario, the primary level association value between the brand page *b* and the subscriber *i* is calculated using the definitions discussed. Thereafter, a threshold ε is chosen and only those subscribers whose association value is greater than this threshold are considered for the set of influencers F.Then, for each influencer node F_i, secondary level association values are computed for each friend S_{ij} using the method described in (Sharma, Alam, Dasgupta, & Saha, 2013). Once the secondary level association values are calculated for all S_{ij}, these values are ranked in descending order and the top M nodes are then considered as the potential customers.

However, this scenario also faces a disadvantage that the potential customer identified may have high secondary level association values, but the subscriber through whom this potential customer is identified has a lower primary association value in comparison to other subscribers who could have better spread word-of-mouth about the brand.Like scenario 1, such customers may not be interested in the products or services of the brand page and may consider the ads posted in their newsfeed as nuisance.

Scenario 2b: Using a composite association value

The shortcomings of the scenario 2a may be resolved by resorting to a composite association value. For this, like in scenario 2a, a threshold value ε is chosen for the primary level association values and the set of influencer nodes F is identified. Thereafter, the association valuesatthe secondary level are computed for all S_{ij} who are part of the friend circle of all chosen influencers F_i. Let A_{bi} denote the primary level association value between brand page b and subscriber i and A_{ij}' denote the secondary level association value between subscriber i and node j. The composite association function between brand page b and node j is defined as

$$C_{bj} = \sqrt{A_{bi}.A_{ij}'}$$

Once the values of the composite association C_{bj} is calculated for all S_{ij} who are in the friend circle of F_i, the composite association values are ranked in descending order andthe top M nodes are chosen as the potential customers. The composite function ensures that the impact of word-of-mouth marketing, since, the geometric mean does not allow the secondary level association value to compensate for the primary level association value.

The Simulation

The purpose of this work is to compare the variation in the increase in reach in each of the three scenarios 1, 2a and 2b. The secondary level association values in scenario 1 and scenario 2a and the composite association value in scenario 2b, in effect measure of, not only how interesting brand page is to a secondary level user, but, also how relevant a particular secondary level user is for targeted marketing. Therefore, for the purpose of this work, this association value is considered as the probability with which a particular secondary level user subscribes to a brand page.

The initial network is simulated and initialized with a brand page, a set of its subscribers and a set of secondary level users. Thereafter, association values are computed in accordance with each of the scenarios 1, 2a and 2b. Next, simulation is used to make users at the secondary level subscribe to the brand page in the probability of their association value. The process is repeated until a level of saturation is reached, when the incremental increase in reach of the brand page is minimal in subsequent iterations. After each iteration the average value of association of all the subscribers with the brand page is calculated and the variation in these values across the three scenarios is illustrated, from which conclusions may be drawn regarding which scenario is best suited for social network based brand promotion.

Implications

Optimization of the newsfeed of its users is an important consideration for all social media network. When the newsfeed is used for advertising, this optimization becomes even more important. Excessive or irrelevant ads can result in the user finding her newsfeed to cluttered and may in the process start ignoring the ads altogether. Even for advertisers and marketers, it is extremely important that their ads on social media get displayed to those users who are interested in their products or services. In addition, social media allows the advertisers and marketers to use the social media as a tool for word-of-mouth marketing.

This work attempts at not only ensuring that the newsfeed shown to a user on a social media website does not create nuisance value, but also, identifying the right kind of users who must be targeted for brand promotion. In addition, in the process of brand promotion word-of-mouth influence is utilized to ensure the maximum returns for the advertiser.

Works Cited

facebook.com. (n.d.). *Advertising on Facebook*. Retrieved from facebook.com: https://www.facebook.com/about/ads/#types

Ha, L. &. (2008). An integrated model of advertising clutter in offline and online media. *International Journal of Advertising, 27*(4), 569-592.

Katona, Z. Z. (2011). Network effects and personal influences: The diffusion of an online social network. *Journal of Marketing Research, 48*(3), 425-443.

Sharma, D., Alam, A., Dasgupta, P., & Saha, D. (2013). A Ranking Algorithm for Online Social Network Search. *Compute.* Vellore: ACM.