Information Propagation in Social Rating Networks

Priyanka Garg
The Chinese University of Hong Kong
priyanka@cse.cuhk.edu.hk

Irwin King
The Chinese University of Hong Kong
king@cse.cuhk.edu.hk

Michael R. Lyu
The Chinese University of Hong Kong
lyu@cse.cuhk.edu.hk

ABSTRACT

The polarity of opinion is a crucial part of information and ignoring the asymmetry between them, can potentially result in an inaccurate estimation of the number of product adoptions and incorrect recommendations. We analyze the propagation patterns of the negative and positive opinions on two real world datasets, Flixster and Epinions, and observe that the presence of negative opinions significantly reduces the number of expressed opinions. To account for the asymmetry between the two kind of opinions, we propose extensions of the two most popular information propagation models, Independent Cascade and Linear Threshold models. The proposed extensions give a tractable influence problem and improves the prediction accuracy of future opinions, by more than 3% on Flixster and 5% on Epinions datasets.

Categories and Subject Descriptors

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

General Terms

Algorithms, Verification

Keywords

Negative Opinions, Information Flow, Social Networks

1. INTRODUCTION

Several probabilistic information flow models like Independent Cascade (IC) and Linear Threshold (LT) models [5] have been developed to mimic the way information spreads in a social network. They attempt to predict the probability of a user to adopt a product given her friends' adoption behavior. The underlying belief is that, as more and more of our friends start believing in something, others also start following them.

However, in real world we not only know the adoption behavior of our friends, but we also know how much they like the product. In general, both the positive and negative opinions emerge as people start adopting products. It is needless to say that the impact of two kind of opinions is asymmetric. While positive opinions promote a product, the negative opinions discourage its adoption. Further, negative opinions usually dominate the positive opinions in shaping one's decision [1]. Even slight hint of product faults, are sometimes sufficient to change our purchase decisions.

The two kind of opinions also differ from each other, in terms of their propagation patterns. Many work in psychology, hypothesize the negative opinions propagate contagiously in the network [8]. Such kind of pattern can be expected in case of shocking news, for example, comments such as "food poisoning from a restaurant food" are likely to be echoed in the network even though a user has not dined there. However, on social rating networks, the negative opinions do not get spread at all. For example, bad reviews about a movie discourage us from watching it, but it is less likely that we pass the negative comments to other friends without watching the movie by ourselves. In fact, we do observe that the presence of negative opinions reduces the number of expressed opinions (either positive or negative) on two real world datasets Flixster[1] and Epinions[2]. Flixster and Epinions are two popular social websites which allow users to rate movies and articles respectively (statistics of the datasets are presented in Section 4). For the two datasets, we categorize products based on the ratio of number of negative ratings and total number of ratings expressed for that product. Then, we plot the average number of opinions in each category against the percentage

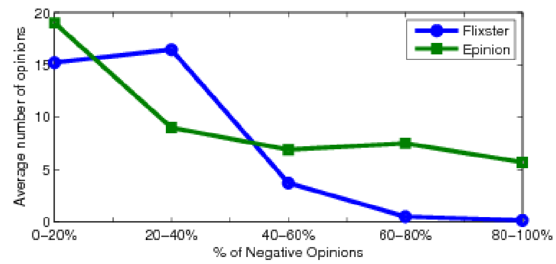

Figure 1: Average number of opinions as percentage of negative opinions varies. For Flixster dataset, y-axis is divided by 100.

[1]http://www.flixster.com/

[2]http://www.epinions.com/

CIKM'12, October 29–November 2, 2012, Maui, HI, USA.

of negative ratings expressed. The plot is shown in Figure 1. We can observe that products with higher percentage of negative ratings, are usually shorter than the one with lower percentage of negative ratings. Further, we observe the probability that a user rates a product (either positively or negatively) depends on her friends' ratings. The probability drops from 10% to 7% for Flixster and from 6% to 1% for Epinions dataset, when more than 50% of a user's friends have negatively rated a product. Both observations underline the asymmetry in the propagation patterns of positive and negative opinions.

Motivated by above observations, we propose polarity sensitive extensions of both IC and LT to model the flow of information in social rating networks. We explicitly consider every opinion expressed by users to be a two step process. First step is the social influence which drives users to consider or not to consider the product. While the second step, describes the expressed opinion given the outcome of first step. The usefulness of the proposed models are demonstrated by predicting the future users' opinions using them. On both Flixster and Epinions datasets, polarity sensitive functions are able to predict the future opinions more accurately. Further, the best accuracy is achieved using the LT based extension.

2. POLARITY-SENSITIVE INFORMATION FLOW MODEL

Let $G = (V, E)$ be a social graph where every node $v \in V$ corresponds to a user and edge $(u, v) \in E$ represents the relationship between node u and v. Nodes in the network publish their opinions related to several products, where every opinion can either have positive polarity or negative polarity. All the opinions expressed related to **one** product are considered as part of one *information cascade*. Node who publishes its opinion is called *active* node. If polarity of its opinion is positive, then it is refereed as *positively active*, while if polarity is negative then it is called a *negatively active* node. All other nodes, who do not publish any opinion are referred to as *inactive* nodes. Thus, in every cascade c, the state s_v^c of every node v is one of the three states: positively active, negatively active and inactive represented by +, - and 0 respectively. The dynamics of every cascade c is considered to unfold in discrete steps, where in every step the influence of nodes activated in previous steps, activate new nodes. We assume that once a node becomes active, it can not change its state.

We consider every activation as a two step process, where in the first step social influence pursues user to consider the product while in the latter step, the user decides its activation state (i.e. to publish its own opinion) based on aggregate social recommendation and its own experience. For the same, we introduce a *hidden* state variable $\tilde{s}_v$ to represent the social influence on node v, where $\tilde{s}_v \in \{+, -, 0\}$ with + indicating positive influence, - indicating negative influence and 0 indicating absence of any influencing opinion. Given the set of already active neighbors $A(v, t)$ of node v at time t,

$$p(s_v|A(v,t)) = \sum_{\tilde{s}_v} p(\tilde{s}_v|A(v,t)).p(s_v|\tilde{s}_v). \quad (1)$$

Based on our observations on the asymmetric propagation pattern of two kind of opinions on social rating networks, we assume the following.

1. If a node gets negative impression about the product, then it does not get activated, that is $p(s_v \in \{+, -\}|\tilde{s}_v = -) = 0$.
2. If the node receives positive product recommendation, then depending upon the product experience it expresses positive or negative opinion. If $q \in [0, 1]$ represents the quality of the product, then $p(s_v = +|\tilde{s}_v = +) = q$ and $p(s_v = -|\tilde{s}_v = +) = (1 - q)$.
3. If there is no social influence then the node v remains inactive, that is $p(s_v = 0|\tilde{s}_v = 0) = 1$.

In the next subsections, we will define various functional forms for $p(\tilde{s}_v|A(v, t))$.

2.1 Polarity-Sensitive IC

Like IC model, we assume every node $u_i \in A(v, t)$ influences node v independently with probability p_{v,u_i,o_i} when u_i is opinionated with opinion o_i. However, considering the completely independent model (considered by IC-N [4]) gives a very complex form for $p(\tilde{s}_v|A(v, t))$; even when we assume that every u_i has same $p_{v,u_i,+}$ and $p_{v,u_i,-}$. Hence, next we propose two simple functions which can be seen as approximation of IC-N.

Independent Activation (IA).
In this function, we first combine all the positively active neighbors to form a super positive node *sp*. While all negatively opinionated neighbors are combined to represent a super negative node *sn*. The probability that the node *sp* influences the node v with positive opinion, is defined as the probability that at least one of the positively active node is able to influence the node v. That is,

$$p_{v,sp,+} = 1 - \prod_{u_i \in A(v,t), o_i = +} (1 - p_{v,u_i,+}).$$

Similarly the probability that the node *sn* influences the node v with negative opinion is

$$p_{v,sn,-} = 1 - \prod_{u_i \in A(v,t), o_i = -} (1 - p_{v,u_i,-}).$$

Then, both the nodes *sp* and *sn* independently try to activate the node v by flipping the biased coins with probability $p_{v,sp,+}$ and $p_{v,sn,-}$ respectively. If both of the coins are head then one of them is chosen at random, and considered as the influencing node. Thus,

$$p(\tilde{s}_v = +|A(v,t)) = p_{v,sp,+}(1 - p_{v,sn,-}) + \frac{1}{2}\, p_{v,sp,+}p_{v,sn,-}. \quad (2)$$

Similarly one can write $p(\tilde{s}_v = -|A(v, t))$. The probability $p(\tilde{s}_v = 0|A(v, t))$ is simply equal to the probability that none of the neighbors are able to influence the node v.

Weight Proportional (WP).
In WP, we set $p(\tilde{s}_v = +|A(v, t))$ to be proportional to the relative weight of positive opinion among the active neighbors.

$$p(\tilde{s}_v = +|A(v,t)) \propto \frac{\sum_{u_i \in A, o_i = +} p_{v,u_i,+}}{\sum_{u_i \in A} p_{v,u_i,o_i}}.$$

To include the possibility that none of the neighbors are able to influence the node v, we multiply the above quantity by the probability that at least one of the node succeeds to activate the node v. Thus,

$$p(\tilde{s}_v = +|A(v,t)) = \frac{\sum_{u_i \in A, o_i = +} p_{v,u_i,+}}{\sum_{u_i \in A} p_{v,u_i,o_i}} \left(1 - \prod_{u_i \in A} (1 - p_{v,u_i,o_i})\right). \quad (3)$$

Similarly one can write $p(\tilde{s}_v = -|A(v,t))$. The probability of not able to influence, is same as that in IA.

2.2 Polarity-Sensitive LT (LT-PS)

Next, we extend the influence function used in LT model to account for the polarity of opinions. To the best of our knowledge, this is the **first** extension of the LT model which considers the polarity of opinions. Like in LT model, each node $v \in V$ is associated with an internal threshold $\theta_v \in [0,1]$ which represents the minimum amount of social influence required for node v to get influenced. Lets assume that $w_{v,u_i,o_i} \in [0,1]$ is weight of influence of node u_i on v when u_i has o_i opinion. Then, we define the social influence on node v as the difference between the sum of influence from positive opinionated nodes and the sum of influence from the negatively opinionated nodes. If the difference is greater than zero then the probability of v getting influenced with positive opinion is defined as

$$p(\tilde{s}_v = +|A(v,t)) = g(b.(\theta_v - f_+(v))), \quad (4)$$

where

$$f_+(v) = \sum_{u_i \in A, o_i = +} w_{v,u_i,+} - \sum_{u_i \in A, o_i = -} w_{v,u_i,-},$$

g is a sigmoid function and is used to keep the probabilities between 0 and 1. The constant b is a hyper-parameter and controls the slope of sigmoid function. We use $b = 20$ for empirical studies. Similarly one can defined $p(\tilde{s}_v = -|A(v,t))$ if the difference between influence from positive opinionated nodes and the negatively opinionated node is less than zero. The probability of not getting with any influence is simply $p(\tilde{s}_v = 0|A(v,t)) = 1 - p(\tilde{s}_v = +|A(v,t)) - p(\tilde{s}_v = -|A(v,t))$.

Note: Unlike IC based influence functions (IC-N, IA, WP), in LT-PS influence function, probability of a node v getting influenced (either with positive or negative opinion) does not always increase as the number of active neighbors increases. For example, lets assume that w_{v,u_i,o_i} (p_{v,u_i,o_i} for IC-N) is same for all the nodes and let it be p. Consider the two scenarios: S_1 when there is only one active neighbor and has positive opinion and S_2: when there are three active neighbors and two of them are positive and one is negative. According to LT-PS, $p(\tilde{s}_v|A(v,t))$ remain same in both scenarios. However for IC-N, the probability of influencing node v will increase from p to $(1-(1-p)^3)$ with $p(\tilde{s}_v = +|A(v,t)) = \frac{2}{3}(1-(1-p)^3)$ and $p(\tilde{s}_v = -|A(v,t)) = \frac{1}{3}(1-(1-p)^3)$. This property of LT-PS function makes it more suitable for modeling social influence for real world data, because in real world the presence of both positive and negative opinions **cancel each others** influence and thereby reduces the overall probability of getting influenced.

3. INFLUENCE ESTIMATION

In order to make predictions using $p(s_v^c|A^c(v,t))$, we need to estimate the values of the pair-wise influence parameters $p_{v,u,+/-}$ ($w_{v,u,+/-}$, θ_v for LT-PS). To learn these values, we use the historical information cascades and maximize the likelihood of observing them. Lets assume that C represents the set of historical information cascade. Then, the log likelihood LL of observing the set of cascades C can be written as the sum of log likelihood of each cascade $c \in C$.

$$LL(C) = \sum_{c \in C} \Big(\sum_{o_v^c \in \{+,-\}} \log\ p(s_v^c = o_v^c | A^c(v, t_v^c)) + \sum_{o_v^c = 0} \log\ p(s_v^c = 0|A^c(v,T)) \Big)$$

Here T is the end of the observation time window of cascades. Our objective is to chose model parameters which maximize the $LL(C)$ and generalize well on the unseen data. Thus, for IC based model, we write our objective function as

$$\max\ LL(C) - \lambda \sum_{v,u} (p_{v,u,+}^2 + p_{v,u,-}^2),$$

and for LT-PS model

$$\max\ LL(C) - \lambda \sum_{v} \theta_v^2 - \lambda \sum_{v,u} (w_{v,u,+}^2 + w_{v,u,-}^2),$$

where λ is a hyper-parameter that controls the amount of regularization. It can be noted that the quality factor q of products gets observed as part of a constant because it is assumed to be constant for each product (cascade). Thus, our objective function is independent of q. Secondly, this big objective function can be minimized, by independently minimizing the objective function for every node $v \in V$; because the parameter set $p_{v,u,+/-}$ (($w_{v,u,+/-}$, θ_v for LT-PS) for every node v, are different from other nodes' parameter set. This makes the inference problem scalable. We minimize each of the sub-problems using the steepest gradient descent method.

4. EXPERIMENTAL EVALUATION

We evaluate the proposed polarity sensitive influence functions in terms of their ability to predict the future activations. In addition to comparing IA, WP and LT-PS with polarity insensitive IC and LT model, we also compare them with other 2 base line methods. Base line 1 and 2 are the global influence functions, where activation of nodes depend on the number of active users (not necessarily neighbors) and their ratings.The base line 1 (BL1) sets the activation probability proportional to total number of active users. If P is the total number of positively active nodes and N is the number of negatively active nodes at time $(t_v^c - 1)$ then, BL1 sets $p(s_v^c \in \{+,-\}) \propto (P+N)/|V|$. The base line 2 (BL2) respects the polarity of opinions. It sets $p(s_v^c = +) \propto (P-N) \cdot (P+N)$ if $P > N$ and $p(s_v^c = -) \propto (N-P) \cdot (N+P)$ if $P < N$. Thus, any improvement over BL1 and BL2, can be attributed to the influence from friends.

Data Collections. To construct the Flixster dataset, we have collected user ratings for all the movies released from Jan, 2005 to Dec, 2010. Only users who have rated at least 50 movies are kept and their friendship network is crawled. Epinions data is taken from [6]. For Epinions, every review article is considered as one product while for Flixster, every movie represents one product. Rating 1-2 are considered as negative and 3-5 as positive rating for Epinions while in

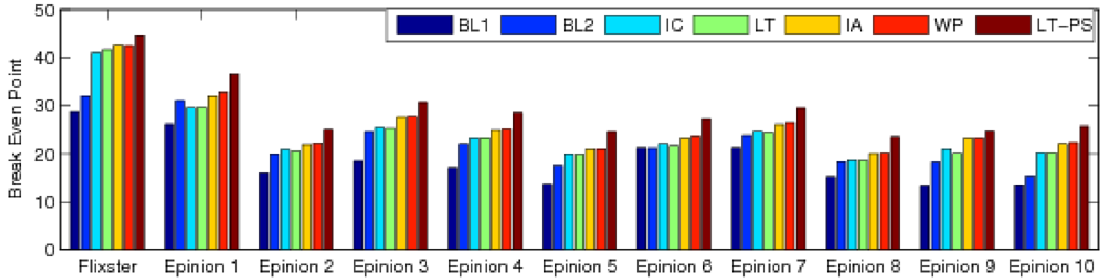

Figure 2: Breakeven point for Flixster and Epinion's 10 most popular categories.

Flixster, 00-25 ratings are considered as negative while 30-50 ratings as positive ratings. Brief statistics of the data is presented in Table 1.

Data set	Users	Edges	Products	Ratings
Flixster	85,209	5,71,505	16,049	10,086,362
Epinions	1,32,000	71,76,671	560,144	13,668,319

Table 1: Data statistics

Experiment Setup. For each dataset, 80% cascades are randomly selected as the train set and test of the cascades are used for constructing the test set. The parameters of IC, LT, IA, WP and LT-PS are learned on the training set. Then, the learned parameters and activation state of one-hop-neighbor nodes are taken as input to predict the activation (either as positive or negative) of a node in test set cascades. Thus, for both polarity-sensitive and polarity-insensitive influence functions, prediction target is same (node's activation). The parameters for IA, WP, LT-PS and LT are learned by our proposed influence estimation method, while for IC we use mosek[3] implementation of the state-of-art method Connie [7]. Connie provides a convex objective function to estimate the IC model parameters and thereby guarantees global optimal solution.

Performance Measure. We assess the quality of prediction accuracy in terms of the break even point on the test set. The break even point is the point at which both precision and recall are equal.

Observations. The results on Flixster dataset and top 10 most popular product categories of Epinions are presented in Figure 2. It can be observe that the break even point is lowest for BL1. By incorporating the polarity of opinions, BL2 improves over BL1 on every dataset. This shows that considering polarity of opinions is very important, even if we just consider global influence. Next we can notice that, both IC and LT outperform BL1 by incorporating the friends' influence. Though, in most cases, IC and LT improves over BL2, in some cases the improvement is not statistically significant; for example in case of Epionion's category 1, 2 and 3. Further, there is not much difference in the performance of IC and LT.

Next, we can note that IA and WP models improve the performance by 1.5% on Flixster and 2% on Epinions as compared to the IC model. It highlights the fact that negative opinions do not spread contiguously in the social rating networks such as Flixster and Epinions. Recall that, in the IA and WP model, a node gets activated only when it is influenced by the positively active neighbor. However, in the IC model, a node can get activated by any (positively or negatively) active neighbor.

Among all the models, LT-PS achieves the best prediction accuracy by accurately modeling the behavior of polarity of opinions. In LT-PS model, when some neighbors of a user are positively active and some negatively active, then the influence of two kind of polarities cancel out each other, and thereby reduces the overall probability of getting activated.

In summary, we observe that accounting for the asymmetry in propagation patterns of two kind of opinions improve the prediction accuracy. Further, the presence of two kind of opinions work against each other and thereby reduces the total probability of activation. This is unlike the competitive information models where the two kind of products compete in the network for adoption [2, 3].

[3]http://www.mosek.com/

5. REFERENCES

[1] R. F. Baumeister, E. Bratslavsky, C. Finkenauer, and K. D. Vohs. Bad is stronger than good. *Review of General Psychology*, 5:323–370, 2001.

[2] S. Bharathi, D. Kempe, and M. Salek. Competitive influence maximization in social networks. In *Internet and Network Economics*, volume 4858 of *Lecture Notes in Computer Science*, pages 306–311. 2007.

[3] A. Borodin, Y. Filmus, and J. Oren. Threshold models for competitive influence in social networks. In *Proceedings of WINE'10*, 2010.

[4] W. Chen, A. Colin, R. Cumming, T. Ke, Z. Liu, D. Rincon, X. Sun, Y. Wang, W. Wei, and Y. Yuan. Influence maximization in social networks when negative opinion may emerge and propagate. In *Proceeding of SDM'11*, 2011.

[5] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of KDD'03*, 2003.

[6] P. Massa and P. Avesani. Trust metrics in recommender systems. In *Computing with Social Trust*, Human-Computer Interaction Series, pages 259–285. Springer London, 2009.

[7] S. A. Myers and J. Leskovec. On the convexity of latent social network inference. In *Proceedings of NIPS'10*, 2010.

[8] P. Rozin and E. B. Royzman. Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 4:296–320, 2001.