

Revenue maximization in social networks through discounting

Mahmoudreza Babaei · Baharan Mirzasoleiman ·
Mahdi Jalili · Mohammad Ali Safari

Received: 27 May 2012/Revised: 26 August 2012/Accepted: 25 September 2012/Published online: 9 October 2012
© Springer-Verlag Wien 2012

Abstract Social networking has become a part of daily life for many individuals across the world. Widespread adoption of various strategies in such networks can be utilized by business corporations as a powerful means for advertising. In this study, we investigated viral marketing strategies in which buyers are influenced by other buyers who already own an item. Since finding an optimal marketing strategy is NP-hard, a simple strategy has been proposed in which giving the item for free to a subset of influential buyers in a network increases the valuation of the other potential buyers for the item. In this study, we considered the more general problem by offering discounts instead of giving the item for free to an initial set of buyers. We introduced three approaches for finding an appropriate discount sequence based on the following iterative idea: In each step, we offer the item to the potential buyers with a discounted price in a way that they all accept the offers and buy the product. Selling the item to the most influential buyers as the opinion leaders increases the willingness of other buyers to pay a higher price. Thus, in the following steps, we can offer the item with a lower discount while still guaranteeing the acceptance of the offers. Furthermore, we investigated two marketing strategies based on local search and hill climbing algorithms. Extensive computational experiments on artificially constructed model networks as well as on a number of real-world networks revealed the effectiveness of the proposed discount-based strategies.

Keywords Pricing · Monetizing social networks · Viral marketing · Revenue maximization · Influence models · Discounting

1 Introduction

Rapid growth of social networks—individuals as nodes and their various forms of acquaintances as connections—has provided a valuable basis for ideas or information to spread quickly among members of a community. The process of network diffusion has been widely studied in previous works (Valente 1996; Kempe et al. 2003; Jackson and Yariv 2006). Adopting new behavior in social networks is a strong motivation for monetizing social networks (Oswald; Seeyle 1992; Domingos and Richardson 2001; Richardson and Domingos 2002; Kempe et al. 2005). While advertising can affect the individuals' decisions by encouraging them to buy a particular product (Walker 2009; Weber 2007), sellers can further increase their profit using intelligent selling strategies. Potential buyers are likely to be affected by decisions of their friends or “world-of-mouth” in a network. For a high-quality product these effects will be positive and can be used as a valuable means for viral advertising and marketing. Thus, a far sighted seller can take advantage of these positive effects or “positive externalities” to make people more likely to buy or even pay a higher price for that product (Hartline et al. 2008).

The idea of marketing a new product using the structural properties of social networks was first proposed by Domingos and Richardson (2001). They suggested that giving the product for free to the most influential nodes in a network would be continued by a cascade of influences in which many other individuals become eager to try the product. In another

M. Babaei · B. Mirzasoleiman · M. Jalili (✉) · M. A. Safari
Department of Computer Engineering, Sharif University
of Technology, Tehran, Iran
e-mail: mjalili@sharif.edu

study, Kempe et al. (2003) discussed the problem of identifying this set of influential nodes in such a network. This set maximizes the subsequent adoption of the good. Recently, Hartline et al. (2008) considered the problem of revenue maximization instead of influence maximization. They studied optimal marketing strategies in which samples of the item are given for free to carefully chosen set of buyers. Then, the item is offered in some sequence to the remaining buyers in a price proportional to the exerted influence on them.

Hartline et al. showed that finding the optimal marketing strategy is NP-Hard. Motivated by its hardness, they identified a simple marketing strategy, called *influence-and-exploit* strategy, as follows. There is a seller and set V of potential buyers. In the influence step, the product is given away for free to a specifically chosen set of influential buyers $S \subseteq V$. In the exploit step, the product is offered in a random sequence to remaining buyers (V/S). In order to maximize the total revenue, (*myopic*) optimal price is offered to each buyer in the sequence. This problem is also not polynomial time computable; however, it can be shown that picking up any random set as the offer sequence in exploit step gives a 1/2 approximation for the second step.

If revenue functions of the buyers are submodular, the optimal influence-and-exploit strategy can be solved in polynomial time (Hartline et al. 2008). Having this assumption and due to the fact that sum of submodular functions is also submodular, the expected revenue as a function of the set S is also submodular. Since the revenue function is not monotone, Hartline et al. (2008) used two algorithms—originally introduced in Feige et al. (2007)—for maximizing *non-monotone* submodular functions: deterministic local search 1/3-approximation algorithm and a randomized local search 0.4-approximation algorithm.

In this study, we consider offering discount instead of offering the item for free to an initial set of buyers, and investigate to how much extent this strategy can boost the profit. The goal is to determine appropriate discounts as well as finding the appropriate sequence of buyers. On one hand, the item should be purchased by the most influential nodes in order to maximize the influence on remaining potential buyers. On the other hand, the item should be purchased by the individuals buying it at the highest price in order to maximize the total revenue. The strategy should be able to make a trade-off between these issues. We introduce three approaches for determining applicable discounts in different model networks as well as a number of real social networks and show that these methods can extensively increase the revenue.

2 Related works

The problem of identifying the most influential nodes in a network falls in the long line of literature on the spread of

social contagions in economy, social sciences, epidemiology, and more recently, in computer science (Agarwal et al. 2012; Domingos and Richardson 2001; Richardson and Domingos 2002; Kempe et al. 2003, 2005; Kitsak et al. 2010; Cha et al. 2011). The main idea is to exploit social network effects to find the target set of k individuals that change in their behavior to be accompanied by a large cascade of influence by which expected number of further adopters of the behavior will become maximized. Typically, the most influential nodes are identified either using network structural properties (Kourtellis et al. 2012; Kempe et al. 2003; Chen et al. 2011), or as a problem in discrete optimization (Goyal et al. 2012; Kempe et al. 2003, 2005). The first approach uses centrality measures such as degree, betweenness or closeness for determining the importance of a node in the adoption process. While the latter tries to iteratively choose an element that provides the largest marginal increase in the function value.

Based on the above discussion, substantial effort has been devoted to find optimal marketing strategies that utilize social influences to maximize adoption of a product in a society (Goyal et al. 2012; Domingos and Richardson 2001; Richardson and Domingos 2002; Kempe et al. 2003; Hartline et al. 2008). The effectiveness of these strategies are often studied using the social contagion models such as linear threshold and independent cascade models or based on the agents' payoff in the context of game theory (Anari et al. 2010; Kempe et al. 2003, 2005). However, these approaches do not consider economic incentives that affect people's decision in buying a product. Economists have shown the importance of customers' behavior on the effectiveness of marketing strategies. For instance, Engel et al. (1968) and Olshavsky and Granbois (1979) showed that customers behave differently when they are involved in buying high-value and low-value products. Other studies relate network externalities to the price that individuals are eager to pay for the product. Saaskilahti (2007) studied the effect of network topology on the monopoly pricing of selling a networked good when market is characterized by buyer's social relations. The importance of network structure has been also shown for the evolution of cooperation, which is the primary driving force behind social welfare (Perc and Szolnoki 2010; Perc et al. 2008; Szolnoki et al. 2008; Perc 2009). The valuation function depends on the type of the product that is to be sold as well as the buying power of the society. If the society is poor, buyers have little valuation for the product and the seller cannot gain considerable benefit from marketing the product. On the other hand, in rich societies, people have higher valuation for the product and marketing strategies are far more effective. Salop (1979) observed that customers' decision for buying a product depends not only on the individual preferences but also on the product price. Finally, Cabral

et al. (1999) proposed models for pricing a durable good by increasing discounted price over time in order to attract low-value buyers. Despite these efforts, there has been relatively little systematic investigation into understanding how individual’s decision depends on the type of product or how social influences affect people’s willingness for buying a good. Recently, Hartline et al. (2008) proposed a model which defines the dependence of adoption on influence and price. In this work, we extend this model for different types of product that are to be sold and investigate marketing strategies that utilize network externalities to find a sequence of decreasing discounted prices that should be offered to the potential buyers in order to maximize the revenue.

3 Preliminaries

In this section, we review influence models and marketing strategies which are applicable for revenue maximization. Consider a network in which an unlimited supply of a good is to be sold to the potential buyers. Let us suppose that producing each unit of the good has no cost for the seller. In this context, potential buyers can be assumed as a set V of all nodes in this network. The valuation of buyer i for the good depends on the other nodes who have already own the good, $v_i: 2^V \rightarrow R^+$, i.e., $v_i(S)$ is the value of the good for buyer i if set S of buyers already own the item. In this setting, each buyer can decide whether or not to buy the product. If the price offered to buyer i is less than or equal to v_i , she buys the good. In general, smaller prices increase the probability of sale.

3.1 Influence model

Influence models might vary in different situations. While influence models are commonly assumed to be monotonically ascending, there are some issues in which the impact of influences is not monotone. In our situation, the buyers who already own the item ($S \subseteq V$) can lead the opinion of other potential buyers. In the presence of positive externalities, the more is the number of buyers in set S , the higher the willingness of the buyers in set V/S will be to buy the item. However, a question arises that whether or not this increasing theme always continues. We claim that it does not continue and there are some issues (for our purpose in the context of marketing) for which the influence function is not monotone. In the following sections, we discuss two types of influence models: the monotone and non-monotone concave graph models. In both models, the buyer’s valuation is a function of the individuals who have already bought the item, i.e., $v_i: 2^V \rightarrow R^+$. In a social network, where each node sees only its neighbors, $v_i(\cdot)$ is a

function of each node’s neighbors in the network, i.e., $v_i = f_i(\sum_{j \in S} w_{ij})$. f_i is the distribution of the valuation of buyer i , S is the set of individuals who have already bought the item and w_{ij} is the influence of node j on node i , i.e., the weight of link e_{ij} . Recall that the seller knows only the distribution of valuation buyer i and not its exact value.

3.1.1 Monotone concave graph model

In this model, the valuation of buyer i is a non-negative monotone concave function $f_i: R^+ \rightarrow R^+$ (Hartline et al. 2008). In other words, v_i increases as more people buy the item. However, as more individuals buy the item, the members of set S increase and the growth of the valuation function decreases. For all $i \in V, S \subseteq V/\{i\}$, we have $v_i(S) = f_i(\sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik})$, and the link weights are derived independently from distribution function F_{ij} . Note that $\sum_{k \in V} w_{ik}$ in the denominator is just a scaling factor for normalizing the influences and does not change the validity of the model. We further discuss this model by an example. Consider a new technology such as fax machine. As more people use this technology, its value for those who have not yet bought a fax increases. As the number of individuals who are using fax increases, they can send and receive documents using this technology to higher number of individuals. Knowing that you can contact with higher number of individuals increases your willingness and even the price you will to pay for buying a fax machine.

3.1.2 Non-Monotone concave graph model

As mentioned, there are some issues for which the valuation function is not monotonically increasing. Consider the case that you want to buy a new cell phone. If a number of your friends have already bought the same cell phone and you have received positive feedbacks from them, your valuation function for buying that particular cell phone increases. However, as more people buy the same cell phone, your tendency to be different will evenly decrease your willingness to buy the same item. This issue may be even clearer in the case of luxury instruments like cloths. In these examples, although growing number of individuals that have bought the same product increases the valuation of the following buyers, the valuation function starts decreasing after a while. Thus, we introduce a new model which captures the properties of these products. In this model, for all $i \in V, S \subseteq V/\{i\}$, we have, $v_i(S) = f_i(\sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik})$, where f_i is a non-negative and non-monotone concave function, $f_i: R^+ \rightarrow R^+$. As before, the link weights are derived independently from distribution function F_{ij} .

3.1.3 More on influence models

In general, monotone concave model can be applied for products in which their applications are the most important factors for people buying them. For example, wireless modems enhance the communications in a population or a powerful antivirus protects your system against damaging malware. For such products, the positive feedbacks that people receive from their trustworthy friends increase the willingness of people for buying the same item.

On the other side of the spectrum, non-monotone concave graph model is best for modeling the human behavior for buying products that their appearance affects people's decision for buying them. For example, for buying a car or a bag, receiving positive feedbacks may raise your willingness for buying the product for a while. However, as more people buy the same item, your willingness to buy the same product decreases, because you may want to express your personality through buying that product and be different from other people. Such non-monotonicity has also been validated by empirical studies. Shokat-Fadaee (2010) studied the effect of influence on retweeting a post in Twitter. It shows that the probability of retweeting a post, given that n of your friends have already retweeted it, increases almost logarithmically and then decreases after a while. Such non-monotonicity has another implication: once sufficient large number of buyers has bought the item, additional sales decrease the willingness of buyers for buying the product. Goyal et al. (2010) used the link structure of online social networks to estimate w_{ij} . Moreover, it is possible to determine the precise form of the functions f_i using similar approach to Backstrom et al. (2006).

3.2 Marketing strategies

In order to increase the total revenue, sellers can take advantage of the influences of buyers on each other. Any marketing strategy has two aspects of pricing and finding the right sequence of offers. Intelligent decisions can be made in both aspects. Through the process of marketing, an item is offered once to each potential buyer in set V . The buyer can accept and pay the offered price or just reject and does not pay the seller. The payment is added to the total revenue for the accepted offers. We assume that each buyer is considered exactly once, thus, the seller can limit the set of potential buyers through the process to those that have not yet bought the item (V/S). We further assume that buyers are only influenced from those that have already bought the item ($S \subseteq V$) and the buyers in the set S have no influence on each other. In order to maximize the revenue, it is assumed that myopic prices are offered to the buyers i.e., the offered price to each buyer is equal to the exerted influences on her which is the maximum value she is willing to pay.

4 Revenue maximization strategies

In the previous section, we mentioned Influence and Exploit (IE) strategy and two approaches (deterministic and randomized local search algorithms) which have been used in Hartline et al. (2008) to find a good approximation for optimal IE strategy. In this section, we show that one can adopt these algorithms to approximate the general revenue maximization problem, in which discounts are offered on the item instead of giving it for free to specifically chosen set $S \subseteq V$. In addition, we show that the greedy hill climbing strategy can be adapted to yield acceptable solution in the same setting.

4.1 Local search

This algorithm was introduced by Feige et al. (2007) for maximizing non-monotone submodular functions. Hartline et al. (2008) used this algorithm to find a set S maximizing the expected revenue of the optimal IE strategy, $g(S)$. Provided that revenue function is non-monotone, they proved that if the buyers' valuation function is submodular, this algorithm finds set S in polynomial time such that the revenue of IE strategy is at least 1/3 approximation of the optimal revenue.

We introduce a deterministic algorithm which is modified in order to find set S which gives an approximation for the revenue maximization problem. It should be noted that, unlike IE strategy, we do not give the item for free to everyone who are added to the set S ; instead, we offer discount. If the buyer accepts the offer, the offered price (instead of myopic price) is added to the revenue. Thus, we try to maximize the revenue while we are trying to maximize the influences. The deterministic algorithm based on a local search approach is as follows:

1. Initialize set $S = \{v\}$, where $v = \operatorname{argmax}_i g(i)$.
2. If neither of the following two steps apply (there is no local improvement), output S .
3. For any buyer $i \in V/S$ that accepts the offered price, $g(S \cup \{i\}) > (1 + \frac{\epsilon}{n^2})g(S)$ then let $S = S \cup \{i\}$.
4. For any buyer $i \in S$, if $g(S/\{i\}) > (1 + \frac{\epsilon}{n^2})g(S)$ then let $S = S/\{i\}$ and go to step 2.

4.2 Greedy hill climbing strategy

Greedy hill climbing algorithm suggested by Nemhauser et al. (1978) is a mathematical optimization technique for finding a local optimum. Kemp et al. used this algorithm in the optimization problem of selecting the most influential nodes in a network. Motivated by the design of viral marketing, a subset of individuals are tried to be convinced to adopt the new product or innovation in order to

maximize the cascade of further adoptions. In the revenue maximization problem, the goal is to maximize the valuation of the buyers or their willingness to pay a higher price in order to increase the revenue. Since the revenue maximization problem is non-monotone, one cannot provide an approximation guarantee for the expected revenue of this algorithm $g(S)$. However, as we discussed previously, buyers' value for the good is a function of the influences from the set of buyers who already own the item. One can use this fact in order to find acceptable solution for revenue maximization. The algorithm based on a hill climbing approach is as follows:

1. Initialize set $S = \emptyset$.
2. Among buyers who accept the offered price choose buyer i so that $i = \arg \max_i g(S \cup \{i\}) - g(S)$.
3. If $g(S \cup \{i\}) \leq g(S)$, output S .
4. $S = S \cup \{i\}$ and go to step 2.

5 Applying discount for revenue maximization

Consider we are interested in finding a set of people S maximizing the revenue function $g(S)$. The items are to be sold to almost all the influential buyers. Furthermore, it is desired to get them buying the item early in the sequence in order to maximize the influence on other potential buyers. Giving the item for free to such influential nodes guarantees the acceptance of the offer. However, can we offer such buyers the item with a discounted price in a way that most of them accept the offer and buy the item? If some buyers who are chosen by a revenue maximization algorithm do not accept the offer, the algorithm should choose a less influential buyer instead. Therefore, the total revenue might decrease with an inappropriate offer sequence. In general, it seems reasonable to offer such influential buyers smaller price to encourage them to buy the item. Later in the sequence, as the buyers' value increases, the item can be offered with smaller discount, and thus, the offered sequence should have an ascending tone. The purpose is to increase the revenue during the process of cascading adoptions in order to find an approximate optimal solution for revenue maximization problem. In the sequel, we introduce three strategies to determine an appropriate discount sequence.

5.1 Discount based on average degree

As discussed above, giving the item for free to influential buyers guarantees the acceptance of offers. It would be interesting to see whether or not one can find an offer sequence which includes positive non-zero elements while

still guaranteeing the acceptance of offers by all influential nodes. We introduce a strategy that uses only the average degree of a network (μ) to determine an appropriate discount sequence. We then show that this approach increases the revenue.

In the first step, we offer the item for free to the buyers until the expected influence on any buyer $i \in V/S$ becomes at least d_i/μ (where d_i is degree of node i and μ is average degree of the network). Suppose that selling the item to k_1 buyers in set S , the expected influence on any potential buyer $i \in V/S$ becomes at least d_i/μ . From this point on, we can offer the item with the price of $f(1/\mu)$ to any buyer in set V/S and guarantee the acceptance of the offers. In the second step, we sell the item with the price of $f(1/\mu)$ to all buyers in set V/S until the expected influence on other potential buyers becomes at least $2d_i/\mu$. Suppose that selling the item to k_2 buyers in the second step has the satisfactory effect. At this point, it is possible to offer the item with the price of $f(2/\mu)$ while guaranteeing the acceptance. Continuing the process, in step j , where $j = 1 \dots n$, $n = \lfloor \mu \rfloor + 1$, we offer the item with the price of $f(j - 1/\mu)$ to all potential buyers. Buying the product by k_j buyers in step j guarantees the acceptance of the offers in step $j + 1$ with the price of $f(j/\mu)$. The process continues to the point that selling the product to k_{n-1} buyers in step $n - 1$ guarantees that everyone in the network accepts the offer with the price of $f(\mu/\mu)$. We claim that in any network $k_1 = k_2 = \dots = k_{n-1} = \lfloor N/\mu \rfloor$. We next prove our claim and show that this strategy provides at least 1/4 approximation of the optimal revenue. Note that this method does not use any information about the structural properties of the network or the revenue maximization algorithm.

Lemma 1 *In a network of size N , for any set S , $|S| \geq k \frac{N}{\mu}$ and all $i \in V/S$, the expected value of v_i is at least $f(\frac{k}{\mu})$.*

Proof Consider real-valued random variable $X(j)$, $j \in S$ as follows:

$$X(j) = \begin{cases} 1 & \text{if } w_{ij} > 0 \\ 0 & \text{if } w_{ij} = 0 \end{cases}$$

where i is an arbitrary node. Expected influence on node i is

$$E\left(\sum_{j \in S \cup \{i\}} [X(j)]\right) = \sum_{j \in S \cup \{i\}} E([X(j)]) \approx \sum_{j \in S \cup \{i\}} \frac{d_i \cdot \bar{F}_{ij}}{N \cdot \bar{F}_{ij}} = k \frac{d_i}{\mu}$$

where \bar{F}_{ij} is the expected value of the distribution from which the link weights are derived.

Thus, the value of buyer j is $f_i(\sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik}) \approx f_i(\frac{k \cdot d_i \cdot \bar{F}_{ij}}{\mu \cdot d_i \cdot \bar{F}_{ij}}) = f_i(\frac{k}{\mu})$.

Lemma 2 Applying discount based on μ on any algorithm, which is applicable to revenue maximization problem, increases the revenue in networks with any structure.

Proof Consider any greedy algorithm applicable to revenue maximization problem. Considering Lemma 1, we guarantee that any node chosen by the algorithm can buy the item with the offered price. As we do not give the item for free to all nodes in set S , the revenue from the greedy algorithm is at least equal to the revenue of the greedy algorithm from the IE strategy.

Hartline et al. showed that the expected revenue from the optimal IE strategy is at least 1/4 of the optimal revenue considering any set of submodular revenue functions R_i . From Lemma 2 we can conclude that discount based on μ gives at least 1/4 approximation of the optimum revenue in any network using local search algorithm.

5.2 Greedy discount approach

Lemma 2 shows that discount base on μ can increase the revenue in networks with any structure. One might use other structural properties to increase the revenue further. It has been shown that many real networks have scale-free degree distribution (Kempe et al. 2003; Newman 2005; Hartline et al. 2008). Scale-free networks have a number of hub nodes with high degrees, while many of the nodes have small degrees. Their degree distribution is power law meaning that the probability of a node with degree d has a power-law relation with d . Any algorithm for revenue maximization should try to get the influential buyers to buy the item early in the sequence in order to increase the value of the following buyers. To do so, the revenue maximization algorithm chooses the nodes of a network in decreasing order of their degree while trying to maximize the influences. Based on this, we introduce a greedy algorithm that extensively improves the revenue maximization algorithms in scale-free networks.

Consider a concave function f as the value function of the buyers. As discussed in previous sections, the valuation of buyer i is a function of buyers who already own the item, $v_i(S) = f_i(\sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik})$. We divide the area under $v_i(\cdot)$ into k regions, and in each step, maximize the number of nodes in some specific interval. The goal is to maximize the number of the nodes, especially influential ones, in some interval and offer them an appropriate discount for the item. This way we guarantee the acceptance of the offers by almost all the candidate buyers chosen by a revenue maximization algorithm. However, the question is how one can determine the appropriate number of regions, k . To this end, one should make a trade-off. The influence regions should not be very small due to the fact that only

few number of high-degree nodes can slightly increase the exerted influence on many potential buyers with small degree and shift them to the next interval. Indeed, many influential buyers still remain in the previous interval. Therefore, each time we are maximizing the number of buyers in one interval, we are indeed interested in including large number of such influential buyers. Recall that influential buyers are those who are highly connected and it is desired such individuals to buy the item early in the sequence. Small number of nodes is not enough to increase the normalized influence on high-degree nodes in V/S , and it can only shift the nodes with smaller degree to the next interval. In addition, the influence regions should not be very large, because a large number of high-degree nodes are needed to shift the buyers into the next interval. Since it is desired to offer a higher price at each step, choosing a large number of nodes reduces the number of steps as well as the revenue. We chose $k = 6$ for our simulations.

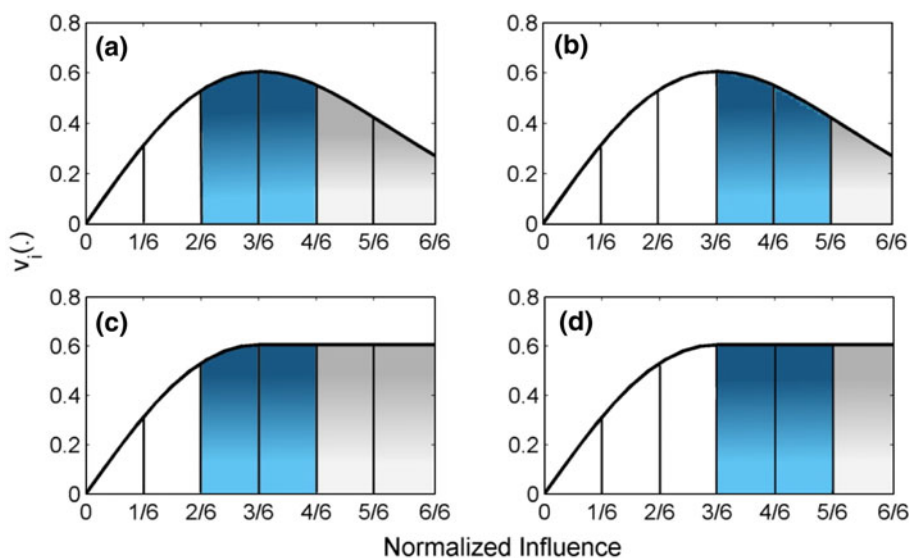
The normalized influences on an arbitrary buyer i can be in the interval $[0, 1]$, i.e., $0 \leq v_i(S) = \sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik} \leq 1$. Therefore, we divide the area under concave value function $v_i(S)$ into $k = 6$ influence region each with width 1/6 as shown in the Fig. 1.

Assume that the revenue maximization algorithm chooses the nodes of the network in decreasing order of their degree. Our proposed greedy algorithm for determining the discount sequence is as follows:

1. Sort the nodes of a network in decreasing order of their degrees in array D .
2. Find $k_1 = \arg \max_k (\sum_{S=D(1:k), i \in V/S | d_i > \mu/2} X(i))$ where
$$X(i) = \begin{cases} 1 & \frac{2}{6} \leq \sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik} < \frac{4}{6} \\ 0 & \text{otherwise} \end{cases}$$
 for all $i \in V/S$
3. Find $k_2 = \arg \max_k (\sum_{S=D(1:k), i \in V/S | d_i > \mu/2} X(i))$ where
$$X(i) = \begin{cases} 1 & \frac{3}{6} \leq \sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik} < \frac{5}{6} \\ 0 & \text{otherwise} \end{cases}$$
 for all $i \in V/S$.
4. Give the item for free to the first k_1 buyers. Next, offer the item with the price of $f(1/6)$ to the following buyers until $k_2 - k_1$ buyers in set $i \in V/S$ accept the offer. Then, offer the item with the price of $f(2/6)$ to the remaining potential buyers. The buyers are chosen by revenue maximization algorithm.

As discussed in this approach, we add the most influential buyers who have not already bought the item to the set S up to the number in the blue interval shown in Fig. 1 becomes maximized. Since the nodes with very small degree hardly exert influence on other nodes in V/S , they are not likely to be chosen by revenue maximization

Fig. 1 $v_i(\cdot)$ as a function of normalized influence for non-monotone concave (top row) and monotone concave (bottom row) influence models. The area under $v_i(\cdot)$ is divided into $k = 6$ influence regions. In each step, the blue interval should be maximized. **a** and **c** show the interval that should be maximized in the first step; **b** and **d** show the interval that should be maximized in the second step (color figure online)



algorithm. We consider the nodes with degree smaller than $\mu/2$ as such nodes. Therefore, we ignore all nodes i with $d_i < \mu/2$ in our greedy algorithm for determining discount sequence. As we give the item for free to k_1 nodes with highest degrees, the number of nodes in the interval $[2/6, 4/6)$ becomes maximized. However, a considerable number of nodes still remain in the influence regions $[1/6, 2/6)$. Indeed, we offer the item with the price of $f(1/6)$ to the following buyers until the number of buyers in set V/S , that exerted influence on them lies in the interval $[3/6, 5/6)$, becomes maximized. At this point, there are still a considerable number of nodes in the influence regions $[2/6, 3/6)$. Therefore, we offer the item with the price of $f(2/6)$ to the following buyers until the end of the marketing process.

In most cases, we can significantly increase the revenue by maximizing the number of buyers in the blue interval (Fig. 1) and try to minimize the number of buyers in the gray interval (Fig. 1). In the monotone model, as the value of buyers in the gray interval can be considerably greater than the offered price, high number of buyers in this interval might decrease the revenue. In the non-monotone model, as the value of buyers in the gray interval may be lower than the offered price, these buyers might never buy the item. If this interval contains a significant number of influential buyers, the total revenue will extensively decrease. The simplest way to have a high number of buyers in the blue interval, while there are few nodes in the gray interval, is to maximize the difference between the number of buyers in the blue and gray intervals. To do so, the second and third step in the above greedy algorithm should be modified as follows:

2. Find $k_1 = \arg \max_k (\sum_{i \in V/S, S=D(1:k)} X(i) - Y(i))$ where

$$X(i) = \begin{cases} 1 & \frac{2}{6} \leq \sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik} < \frac{4}{6} \\ 0 & \text{otherwise} \end{cases} \quad \text{and}$$

$$Y(i) = \begin{cases} 1 & \frac{4}{6} \leq \sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik} < \frac{6}{6} \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } i \in V/S.$$

3. Find $k_2 = \arg \max_k (\sum_{i \in V/S, S=D(1:k)} X(i) - Y(i))$ where

$$X(i) = \begin{cases} 1 & \frac{3}{6} \leq \sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik} < \frac{5}{6} \\ 0 & \text{otherwise} \end{cases} \quad \text{and}$$

$$Y(i) = \begin{cases} 1 & \frac{5}{6} \leq \sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik} < \frac{6}{6} \\ 0 & \text{otherwise} \end{cases} \quad \text{for all } i \in V/S.$$

5.2.1 Discount based on standard deviation of the degree distribution

The greedy algorithm for determining discount sequence should have information about the existence of the links in the network. But, what if we do not have this information? Can we still use the power-law property to get better result than discounting based on μ ? It is advantageous to get influential buyers early in the sequence in order to increase the value of the following buyers. Thus, it makes sense to offer the item to the most influential buyers for free to guarantee that they accept the offer and buy it. The most influential nodes can be identified considering their degree, since the buyers with high degree can trigger many other buyers to buy the item. If we just know the degree distribution of a network, we can extract some information about the variability of its degrees. Standard deviation (σ) is a widely used measure for this purpose. In scale-free networks with power-law exponent of $2 \leq \gamma \leq 3$, which is the case for many real networks, we can consider all nodes with degree larger than $\mu + \sigma$ as high-degree nodes, i.e., the most influential nodes. Suppose there are k_1 nodes satisfying this condition. Giving the item for free to the k_1

most influential buyers in a network, many other buyers will pay us for the item. Among them, we want the remaining influential buyers to buy the item to further increase the value of the following buyers. In general, all nodes i for which $\mu + \sigma > d_i > \mu$ can be simply identified as highly connected, and thus, influential nodes in a network. Assuming the number of such nodes as k_2 , one can offer the item with the price of $f(1/6)$ to the following buyers in set V/S until k_2 buyers accept the offer. This will further increase the value of the remaining buyers, and thus, from this point on, we offer the item with the price of $f(2/6)$ to all remaining buyers in set V/S . Simulation results showed that this approach considerably outperforms those based on average degree in scale-free networks.

6 Experiments

In this section, we investigate the revenue maximization in artificially constructed model networks as well as several real networks.

6.1 Network data

6.1.1 Model networks

While available models for construction of scale-free networks can capture many structural properties of real networks, including heavy tail in-degree and out-degree distributions, communities and small world phenomenon (Barabási and Albert 1999), they cannot model the evolution of real networks over time. Forest-fire model suggested by Leskovec et al. (2007) generates networks with densification power law and shrinking diameter properties. These networks become denser through time while having an increasing average degree. Moreover, as network grows, its diameter decreases. These properties have been observed in a number of real networks including the Internet, citation, affiliation and patents networks (Leskovec et al. 2007). A network based on forest-fire model is constructed through a recursive process as follows. Each node joining the network chooses a random ambassador w . Then, it selects outlinks of w with probability p and its inlinks with probability p_b and forms outlinks to the other end of these links.

It has been shown that many social and biological networks have modular structure (Girvan and Newman 2002; Newman 2006; Zaidi 2012). These networks are composed of several groups in which nodes are highly connected within the groups, while there are only a small number of links between the groups. As our second network model, we considered modular forest-fire networks constructed through an algorithm as follows. First n isolated modules

each with forest-fire structure are built. Then, with probability p each, intra-modular link is disconnected and a connection is created between two random nodes from two randomly chosen modules (Babaei et al. 2011).

6.1.2 Real networks

Although model networks can capture many structural properties of real networks, in many cases, real networks have a more complex structure. As online social networks provide a convenient setting for applying the revenue maximization algorithms, we also considered a number of real social networks and compared the performance of applying discount in these networks.

Facebook-like social network This is an undirected and unweighted network representing an online community of students at the university of California, Irvine. In construction of this network, students are considered as nodes and two students considered to be connected by a link when there is at least a sent or received message between them. The network has 1,899 nodes and 20,296 edges (Opsahl and Panzarasa 2009).

Newman's scientific collaboration network This is the co-authorship network based on preprints posted to Condense Matter section of arXiv E-Print collected between 1995 and 1999 (Newman 2001). The set of nodes are defined by the set of papers and edges represent citation. This network contains 16,726 nodes and 47,594 edges.

Wikipedia vote network This network is defined by votes for Wikipedia admin candidates. A small fraction of Wikipedia users are administrators. Users in Wikipedia issue a request for adminship and Wikipedia community decides whom to promote to adminship with votes. Wikipedia users are considered as nodes in the network and a link present between two nodes represents that one of them votes another. The network consists of 7,115 nodes and 103,689 edges (Leskovec et al. 2010a, b).

High-energy physics theory citation network We analyzed the citation network containing the 27,770 papers. A link between two papers indicates that one of them cites another. The network used in this work is from the e-print arXiv and cover all the citations within 27,770 papers. The network is composed of 27,770 nodes and 352,807 edges (Gehrke et al. 2003; Leskovec et al. 2007).

6.2 Influence models

We discussed two types of influence models, monotone and non-monotone concave models. In the case of non-monotone concave model, there are some examples (such as buying cloths) for which the valuation function of the buyers starts decreasing while just a few of their friends buy the same item. In some other examples (such as buying

a laptop), the buyers' valuation increases as they receive more positive feedbacks from their friends who already own the item. Indeed, valuation function remains increasing for a long time and then becomes decreasing. Averaging over all possibilities, we choose a concave function which reaches its peak value at 1/2 as our non-monotone valuation function.

Following the above discussion, we chose a Rayleigh probability density function $v_i(\cdot) = f(y|b) = \frac{y}{b^2} e^{-\frac{y^2}{2b^2}}$, $b = 1$, $y = 2x$ as our non-monotone concave influence model. We considered $y = 2x$ to maximize the value of a buyer when the normalized influence on her is 1/2, i.e., $\sum_{j \in S \cup \{i\}} w_{ij} / \sum_{k \in V} w_{ik} = 1/2$. Note that Rayleigh function with parameter $b = 1$ is concave in the interval $[0, 2]$. For the second model, we made $v_i(\cdot)$ constant after $y = 1$ in order to have a monotone concave value function. Choosing this monotone concave function makes the result of applying the revenue maximization algorithms on these two models comparable (Fig. 2).

Certainly, assuming a function which takes values greater than $y = 0.6$ as valuation function of the buyers considerably improves the revenue of the discount-based marketing strategies. It is obvious that considering functions with higher concavity as monotone concave valuation function further boosts the revenue from applying discount methods on marketing strategies. For non-monotone concave influence model, the longer it takes to reach the peak value, the higher will be the revenue of the discount-based marketing strategies.

6.3 Algorithms and implementation

We investigated the effects of applying the discount methods introduced in the previous sections on local search and greedy hill climbing algorithms. We also considered, as a baseline, the IE strategy which gives the item for free to all buyers is set $S \subseteq V$. In all of these cases, it has been assumed that the seller does not know the exact value of each buyer, but instead, she knows the distribution from which its values are drawn. When we do not offer the item for free to all buyers who are going to be added to the set

$S \subseteq V$, some of the buyers chosen by the algorithms may not accept our offers and the algorithm should choose a less influential node instead. To have satisfactory approximation of the performance of our algorithms, we also compared them with the case in which the seller has exact information about the value function of the buyers. With this information, in any iteration, local search and greedy hill climbing algorithms can choose the best buyer who pays her myopic price and maximizes the revenue. It should be noted that in this case the revenue maximization problem becomes monotone and greedy hill climbing algorithm provides a $(1 - 1/e)$ approximation.

We discussed that the link weights are derived independently from distribution function F_{ij} . Since we do not have information about the exact link weights in the network (in both implementation of deterministic local search and greedy hill climbing algorithms as well as in the greedy algorithm for determining the appropriate discount sequence), we should estimate the link weights w_{ij} by repeated sampling from distribution F_{ij} . In our implementation we considered F_{ij} as a uniform distribution on the interval $[0, 2]$.

6.4 Results

Figures 3 and 4 show the revenue as the number of buyers in set S increases based on deterministic local search and greedy hill climbing algorithms for forest-fire and modular forest-fire models, respectively. The networks are with 1,000 nodes, $p = 0.37$, $p_b = 0.32$ (Fig. 3) and modular networks with three modules with 200, 300 and 500 nodes (Fig. 4). Each module was constructed with $p = 0.37$ and $p_b = 0.32$ and with inter-modular rewiring probability of $p = 0.01$. For each algorithm, the results are plotted for monotone and non-monotone influence models. On average, in these networks, the greedy discount approach outperformed the hill climbing-based IE strategy by over 21 % in the monotone and over 12 % in the non-monotone influence model. For local search algorithm, this improvement was about 22 % in the monotone and over 15 % in the non-monotone influence model. For the μ -discount approach, the result of both greedy hill climbing

Fig. 2 $v_i(\cdot)$ as a function of $y = 2x$ normalized influence for **a** non-monotone concave and **b** monotone concave influence models

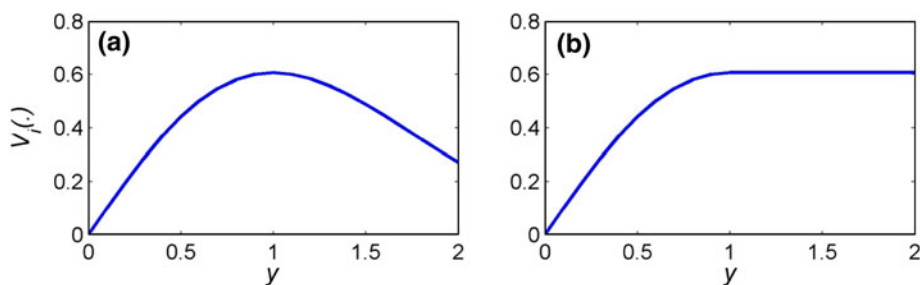


Fig. 3 Revenue from the monotone concave influence model (top row) and the non-monotone concave influence model (bottom row) as a function of the maximum number of buyers allowed in set S , for the forest-fire network with 1,000 nodes, $p = 0.37$ and $p_b = 0.32$. **a** and **c** show the results of greedy hill climbing algorithm; **b** and **d** show the results of deterministic local search algorithm. The blue, green, red and cyan lines correspond to the revenue from IE strategy, discounting strategy based on μ , greedy discount strategy, and the case in which the seller has exact information about the value of the buyers, respectively. The results are averaged over 10 realizations (color figure online)

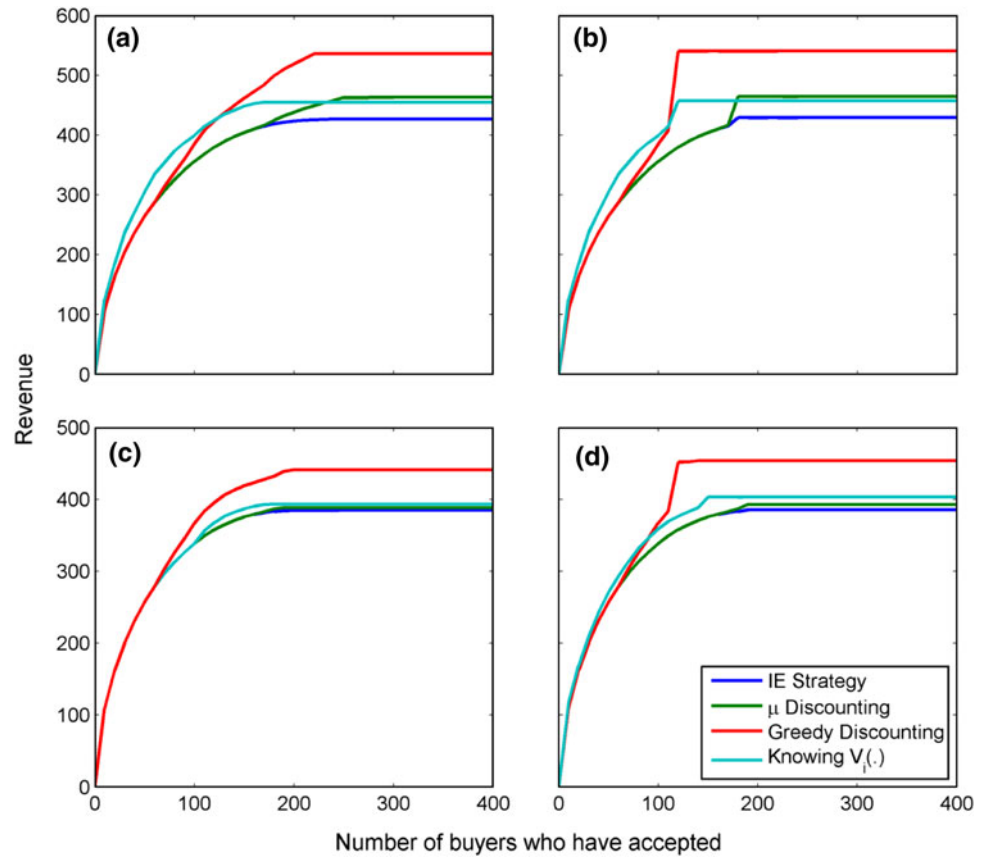


Fig. 4 Revenue from the monotone concave influence model (top row) and the non-monotone concave influence model (bottom row) as a function of the maximum number of buyers allowed in set S , for the modular forest-fire network that has three modules with 200, 300 and 500 nodes. Each module is constructed with $p = 0.37$ and $p_b = 0.32$ and with inter-modular rewiring probability $p = 0.01$. Other descriptions are as in Fig. 3

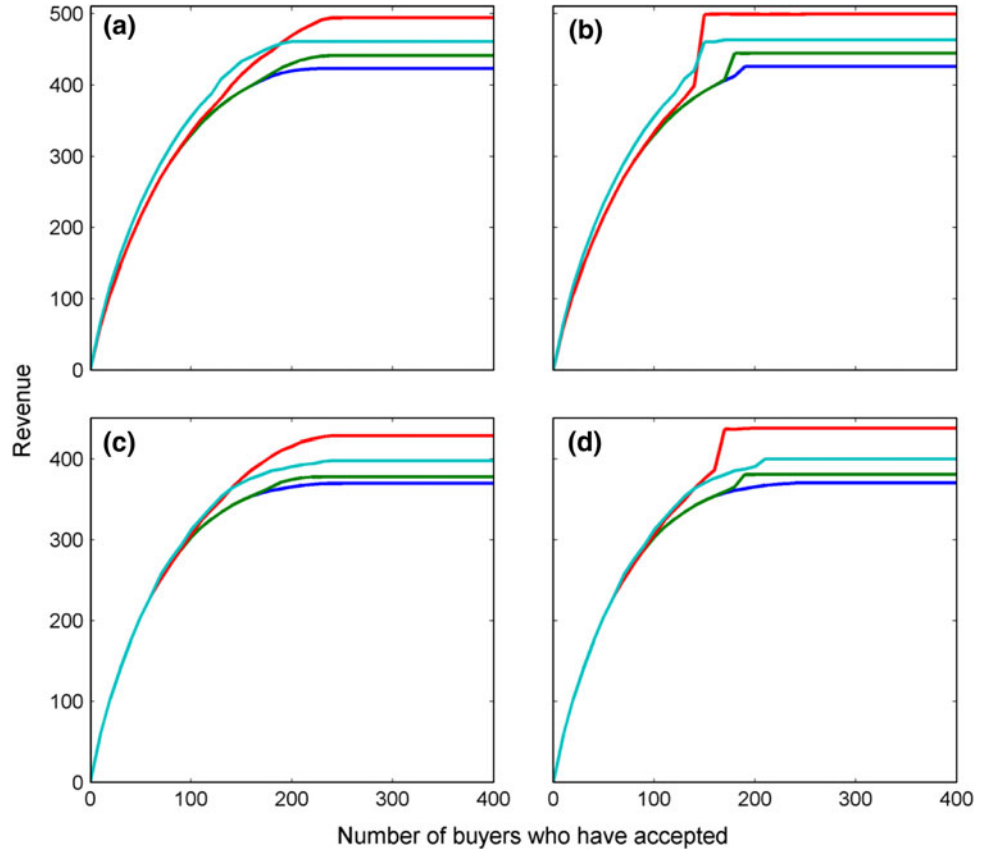
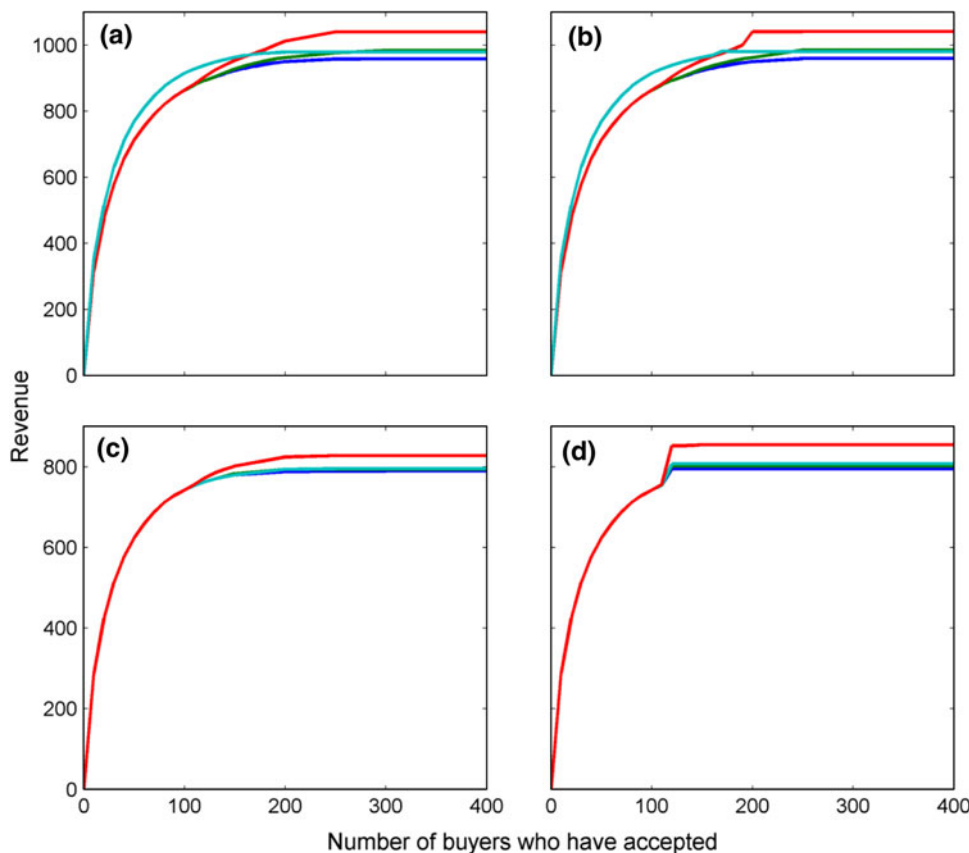


Fig. 5 Revenue from the monotone concave influence model (*top row*) and the non-monotone concave influence model (*bottom row*) as a function of the maximum number of buyers allowed in set S , for the Facebook-like social network with 1,899 nodes and 20,296 edges. Other descriptions are as in Fig. 3



and local search algorithms was improved about 7 % in monotone and over 2 % in non-monotone model. Note that, knowing the exact information about the value function of the buyers, on average, improved the result about 4 % in all cases. As shown in these figures, in these network types, the revenue from the greedy discounting method was considerably higher than the case where the seller knows the exact valuation of each buyer. This can be explained considering the fact that some influential buyers chosen by the algorithm might not accept the offered price determined by the discount method, which in this case, the algorithm should choose a less influential buyer who accepts the offer. This buyer can exert positive influence on some influential buyer who has rejected the offer in previous iterations. As the valuation of this buyer increases, she accepts the offer and increases the influences on the buyers in set V/S . Furthermore, she pays a higher price than before and further increases the revenue. A similar pattern was observed for the case of monotone influence models using μ -discounting approach. It should be noted that this can only happen in the case of using an appropriate discount sequence; otherwise, rejection of the offers by the influential buyers can considerably decrease the revenue compared to the IE strategy.

Figures 5, 6, 7 and 8 show the result of the same experiments on the real networks including Facebook-like

social network, wiki-vote, Newman scientific collaboration, and high-energy physics theory citation networks, respectively. In all these networks, similar qualitative behavior was observed from the discounting strategies. On average, in these networks, the greedy discount approach improved the revenue of the greedy hill climbing algorithm by about 12 % in the monotone and over 8 % in the non-monotone influence model. For the local search algorithm, this approach improved the revenue by about 12 % in the monotone and about 10 % in the non-monotone influence model. On the other hand, knowing exact information about the value function of the buyers, on average, improved the result about 5 % in all cases. The fact that having exact information about the buyers' valuation guarantees a $(1 - 1/e)$ approximation for the optimal revenue shows high performance of our greedy algorithm for determining discount sequence.

7 Conclusion

We investigated the problem of revenue maximization in social networks. An intelligent seller can take advantage of network diffusion as a powerful means for advertisement to further increase the revenue of marketing. For a high-quality good, the positive feedback that individuals receive

Fig. 6 Revenue from the monotone concave influence model (*top row*) and the non-monotone concave influence model (*bottom row*) as a function of the maximum number of buyers allowed in set S , for the wiki-vote network with 7,115 nodes and 103,689 edges. Other descriptions are as in Fig. 3

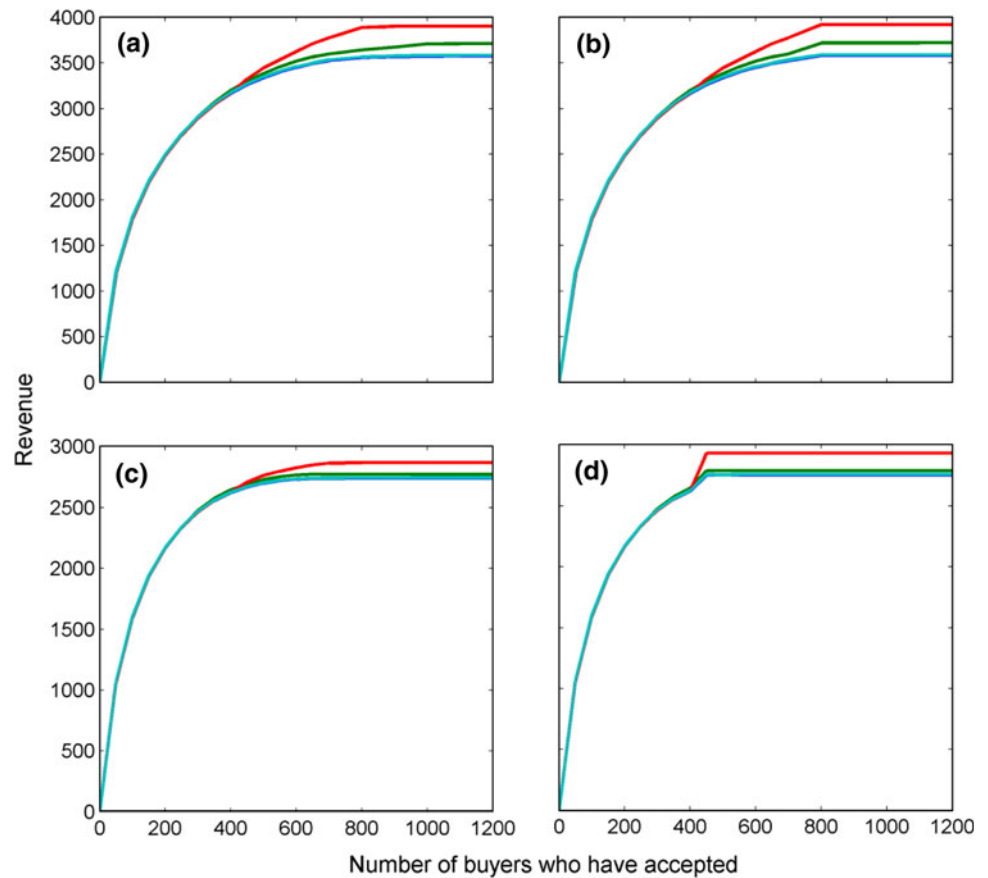


Fig. 7 Revenue from the monotone concave influence model (*top row*) and the non-monotone concave influence model (*bottom row*) as a function of the maximum number of buyers allowed in set S , for the Newman scientific collaboration network with 16,726 nodes and 47,594 edges. Other descriptions are as in Fig. 3

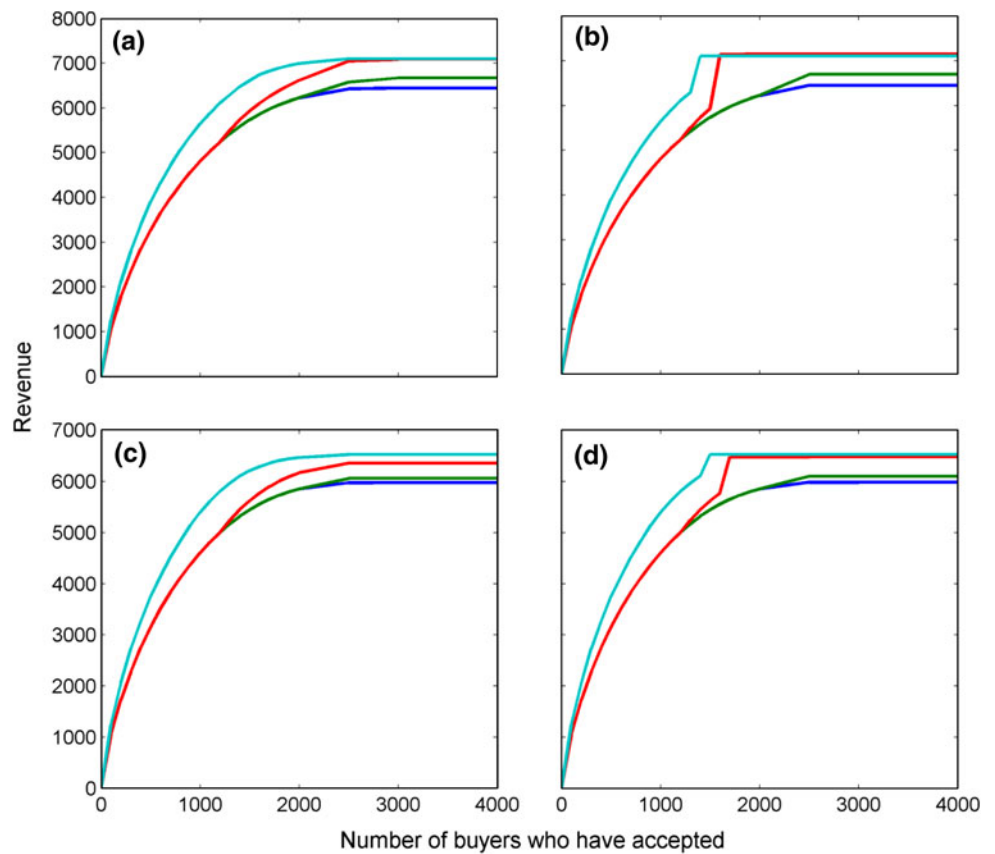
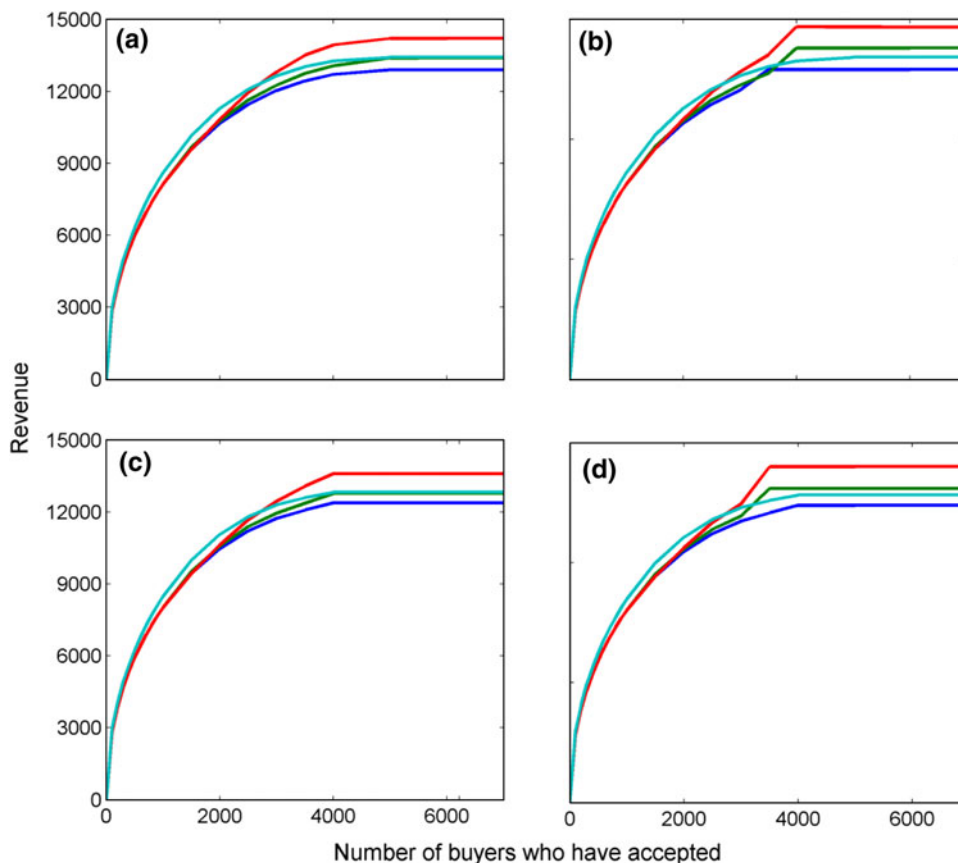


Fig. 8 Revenue from the monotone concave influence model (*top row*) and the non-monotone concave influence model (*bottom row*) as a function of the maximum number of buyers allowed in set S , for the high-energy physics theory citation network with 27,770 nodes and 352,807 edges. Other descriptions are as in Fig. 3



from their friends increases their willingness for buying that product. Motivated by the NP-hardness of finding an optimal marketing strategy, optimal influence-and-exploit strategy has been introduced in this context. Assuming that the seller has only distributional information about the buyers' valuation, this strategy tries to maximize the revenue of marketing by giving the item for free to a specifically chosen set of buyers in order to maximize the influence on other potential buyers in the network. The goal of this article was to generalize this idea by offering discounts instead of giving the item for free. Finding an appropriate discount sequence would enable us to find a good approximation for the general problem of optimal marketing strategy. It is advantageous to get the influential buyers to buy the item in order to maximize the influence on other buyers. Small offers increase the likelihood of a sale by such buyers; however, offering high discount to all the buyers decreases the revenue as well.

If we offer the item for free to the most influential buyers early in the sequence, we increase the value other buyers have for the item. Then, we can offer the item with higher price to some other influential buyers to further increase the value of the remaining buyers. As buyers' value increases, we can offer the item with lower discount and increase the revenue. Based on this idea, we introduced

three approaches for finding a convenient offer sequence: (i) discounting base on average degree which improves the result of the revenue maximization algorithms in networks with any structure; (ii) greedy discount approach which significantly increases the revenue in scale-free networks but should have information about the existence of the links; and (iii) discounting based on standard deviation of the degree distribution which approximates the greedy discount strategy using just the average and standard deviation of the degree distribution. Moreover, we studied two marketing strategies based on deterministic local search and greedy hill climbing strategy. Influence models are usually considered to be monotone. In addition to the monotone concave influence model, we introduced a non-monotone concave model for the influences and tested our strategies on monotone and non-monotone concave influence models. Extensive simulations on model networks as well as several real social networks showed outperformance of our strategy to the traditional algorithms. On average, the greedy discount strategy improved the performance of the local search algorithm by about 15 % in monotone and over 11 % in non-monotone model. In the case of the greedy hill climbing algorithm, the improvement was about 15 % in monotone model and over 9 % in non-monotone model. The proposed strategies are

convenient for use in real world as well as online social networks. As future work, it would be interesting to implement these algorithms in real social networks to validate them and study how individuals respond to these discounts.

Acknowledgments We would like to thank Dr. S. V. Mirrokni for his support and insightful comments throughout preparing this manuscript.

References

- Agarwal N, Liu H et al (2012) Modeling blogger influence in a community. *Soc Netw Anal Min* 2(2):139–162
- Anari N, Ehsani S et al (2010) Equilibrium pricing with positive externalities. *Internet Netw Econ* 6484:424–431
- Babaei M, Ghassemieh H et al (2011) Cascading failure tolerance of modular small-world networks. *IEEE Trans Circuits Syst II Express Briefs* 58(8):1–5
- Backstrom L, Huttenlocher D et al (2006) Group formation in large social networks: membership, growth, and evolution. ACM, New York
- Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509
- Cabral L, Salant DJ et al (1999) Monopoly pricing with network externalities. *Int J Ind Organ* 17(2):199–214
- Cha M, Perez JAN et al (2011) The spread of media content through blogs. *Soc Netw Anal Min* 1–16. doi:10.1007/s13278-011-0040-x
- Chen D, Lu L et al (2011) Identifying influential nodes in complex networks. *Phys A Stat Mech Appl* 391(4):1777–1787
- Domingos P, Richardson M (2001) Mining the network value of customers. In: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, pp 57–66
- Engel JF, Kollat DT et al (1968) Consumer behavior. Holt, Rinehart and Winston, New York
- Feige U, Mirrokni VS et al (2007). Maximizing non-monotone submodular functions. In: Proceedings of the 48th annual IEEE/symposium on Foundations of computer science, pp 461–471
- Gehrke J, Ginsparg P et al (2003) Overview of the 2003 KDD Cup. *ACM SIGKDD Explor Newsl* 5(2):149–151
- Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* 99(12):7821
- Goyal A, Bonchi F et al (2012) On minimizing budget and time in influence propagation over social networks. *Soc Netw Anal Min* 1–14. doi:10.1007/s13278-012-0062-z
- Goyal A, Bonchi F et al (2010) Learning influence probabilities in social networks. In: Proceedings of the third ACM international conference on web search and data mining, New York, pp 241–250
- Hartline J, Mirrokni V et al (2008) Optimal marketing strategies over social networks. In: Proceedings of the 17th international conference on World Wide Web, pp 189–198
- Jackson MO, Yariv L (2006) Diffusion on social networks. *Economie publique/Public Econ* 1(16):69–82
- Kempe D, Kleinberg J et al (2003) Maximizing the spread of influence through a social network. In: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, pp 137–146
- Kempe D, Kleinberg J et al (2005) Influential nodes in a diffusion model for social networks. *Automata Lang Program* 3580(99)
- Kitsak M, Gallos LK et al (2010) Identification of influential spreaders in complex networks. *Nat Phys* 6(11):888–893
- Kourtellis N, Alahakoon T et al (2012) Identifying high betweenness centrality nodes in large social networks. *Soc Netw Anal Min* 1–16. doi:10.1007/s13278-012-0076-6
- Leskovec J, Huttenlocher D et al (2010a) Predicting positive and negative links in online social networks. In: Proceedings of the 19th international conference on World Wide Web, pp 641–650
- Leskovec J, Huttenlocher D et al (2010b) Signed networks in social media. In: Proceedings of the 28th international conference on Human factors in computing systems, pp 1361–1370
- Leskovec J, Kleinberg J et al (2007) Graph evolution: densification and shrinking diameters. *ACM Trans Knowl Discov Data (TKDD)* 1(1):1–40
- Nemhauser GL, Wolsey LA et al (1978) An analysis of approximations for maximizing submodular set functions—I. *Math Program* 14(1):265–294
- Newman MEJ (2001) The structure of scientific collaboration networks. *Proc Natl Acad Sci USA* 98(2):404
- Newman MEJ (2005) Power laws, Pareto distributions and Zipf’s law. *Contemp Phys* 46(5):323–351
- Newman MEJ (2006) Modularity and community structure in networks. *Proc Natl Acad Sci USA* 103(23):8577
- Olshavsky RW, Granbois DH (1979) Consumer decision making—fact or fiction? *J Consumer Res* 6(2):93–100
- Opsahl T, Panzarasa P (2009) Clustering in weighted networks. *Social networks* 31(2):155–163
- Oswald E (2006). <http://www.betanews.com/article/Google-Buys-MySpace-Ads-for-900m/1155050350>
- Perc M (2009) Evolution of cooperation on scale-free networks subject to error and attack. *New J Phys* 11:033027
- Perc M, Szolnoki A (2010) Coevolutionary games—a mini review. *BioSystems* 99(2):109–125
- Perc M, Szolnoki A et al (2008) Restricted connections among distinguished players support cooperation. *Phys Rev E* 78(6):066101
- Richardson M, Domingos P (2002) Mining knowledge-sharing sites for viral marketing. In: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, pp 61–70
- Saaskilampi P (2007) Monopoly pricing of social goods. MPRA Paper 3526, University Library of Munich, Germany
- Salop SC (1979) Monopolistic competition with outside goods. *Bell J Econ* 10(1):141–156
- Seelye KQ (1992). <http://www.nytimes.com/2006/08/23/technology/23soft.html>
- Shokat-Fadaee S (2010) Analysis of effective algorithms in social networks. Master of Science Software Engineering, Sharif University of Technology, Computer Engineering Department
- Szolnoki A, Perc M et al (2008) Making new connections towards cooperation in the prisoner’s dilemma game. *EPL (Europhysics Letters)* 84:50007
- Valente TW (1996) Social network thresholds in the diffusion of innovations. *Social Networks* 18(1):69–89
- Walker R (2009). <http://www.slate.com/id/1006264>
- Weber T (2007). <http://news.bbc.co.uk/1/hi/business/6305957>
- Zaidi F (2012) Small world networks and clustered small world networks with random connectivity. *Soc Netw Anal Min* 1–13. doi:10.1007/s13278-012-0052-1