

Optimal Marketing Strategies over Social Networks

Jason Hartline*
Electrical Engineering and
Computer Science
Northwestern University
Evanston, IL 60208
hartline@eecs.
northwestern.edu

Vahab S. Mirrokni
Microsoft Research
One Microsoft Way
Redmond, WA 98052
mirrokni@theory.
csail.mit.edu

Mukund Sundararajan†
Stanford University
mukunds@cs.stanford.edu

ABSTRACT

We discuss the use of social networks in implementing viral marketing strategies. While *influence maximization* has been studied in this context (see Chapter 24 of [10]), we study *revenue maximization*, arguably, a more natural objective. In our model, a buyer's decision to buy an item is influenced by the set of other buyers that own the item and the price at which the item is offered.

We focus on algorithmic question of finding revenue maximizing marketing strategies. When the buyers are completely symmetric, we can find the optimal marketing strategy in polynomial time. In the general case, motivated by hardness results, we investigate approximation algorithms for this problem. We identify a family of strategies called *influence-and-exploit strategies* that are based on the following idea: Initially *influence* the population by giving the item for free to carefully a chosen set of buyers. Then extract revenue from the remaining buyers using a 'greedy' pricing strategy. We first argue why such strategies are reasonable and then show how to use recently developed set-function maximization techniques to find the right set of buyers to influence.

Categories and Subject Descriptors

F.2 [Theory of Computation]: Analysis of Algorithms and Problem Complexity; J.4 [Computer Applications]: Social and Behavioral Sciences—*Economics*

General Terms

Algorithm, Theory, and Economics.

Keywords

Pricing, Monetizing Social Networks, Marketing, and Sub-modular Maximization.

*Work done while author was at Microsoft Research, Silicon Valley.

†Work done while author was an intern at Microsoft Research.

1. INTRODUCTION

The proliferation of social-networks on the Internet, has allowed companies to collect information about social-network users and their social relationships. Social networks like MySpace, Facebook, and Orkut allow us to determine who is acquainted with whom, how frequently they interact online, what interests they have in common, etc. Users are spending increasing amounts of time on social network websites. For instance, a recent survey [11] that ranks websites based on 'average time spent by a user', identifies MySpace and Facebook among the top 10 websites.

There have been several efforts to monetize social networks [15, 17]. While most proposals are based on advertising [20], the focus of this paper is to monetize social networks via the implementation of intelligent selling strategies. Consider a seller interested in selling a specific good or service. A sale to one buyer often has an impact on other potential buyers. Such an effect is called the *externality* of the transaction. Externalities that induce further sales and revenue for the seller are called *positive externalities*. Here are examples of how such positive externalities arise:

- Information about goods often propagates by *word of mouth*. For instance, we may become aware of, and even be influenced to buy, a specific good or service because our friends own them. When our friends own a copy of a good, we can assess its quality before we make a decision to buy. With high quality goods, this influences us to buy the good and even increases how much we are willing to pay for it.
- Sometimes goods have features that explicitly aid social-networking. For instance, Microsoft music player, the *Zune*, has a music sharing feature that allows it to wirelessly exchange music with other Zunes. Clearly, the value of such a feature is a function of the number of acquaintances who also own the good.

A far sighted seller can take advantage of the existence of positive externalities to increase its revenue. For instance, in order to influence many buyers to buy the good, the seller could initially offer some popular buyers the good for free. Indeed such selling techniques are already employed in practice. TiVo, a company which makes digital video recorders, initially gave away its digital video recorder for free to a select few video enthusiasts [19]. Such promotions may be an effective way to create a buzz about the product.

The basic idea of giving away the item for free can be generalized in a couple of ways: First, rather than offering the item for free, sellers could offer discounts. There is a trade-off: larger discounts decrease the revenue earned from the transaction while increasing the likelihood of a sale and the influence on future buyers. *How large should the discounts be?* Second, the sequence in which sales happen has an impact on the effect of externalities. Influence is generally not symmetric. Often popular, well-connected users wield more influence. Clearly, we would like sales that have the potential to cause further sales to occur earlier. *In what sequence should the selling happen?* The goal of this paper is to explore marketing strategies that optimize a seller’s revenue.

Though the model and the algorithms that we propose are not specific to online social networks, the algorithms may be convenient to implement in such settings. First, in such settings, it is easy to collect information about the influence of buyers on each other; links between user profiles may be reasonable (though they are not likely to be completely accurate) indicators of the influence that the owners of user profiles have on each other. Second, sellers can easily target social network users with specific offers.

1.1 Our Contributions

We investigate marketing strategies that maximize revenue from the sale of digital goods, goods where the cost of producing a copy of the good is zero. There is a seller and set V of potential buyers. We assume that a buyer’s decision to buy an item is dependent on other buyers owning the item *and* the price offered to the buyer; for buyer i , the value of the buyer for the good is defined by a set function $v_i : 2^V \rightarrow R^+$. These functions model the influence that buyers have on other buyers. We assume that though the seller does not know the value functions, but instead has distributional information about them. In general, smaller prices increase the probability of sale. (See Section 2 for details.)

We consider marketing strategies, where the seller considers buyers in some sequence and offers each buyer a price. When the buyer accepts the offer, the seller earns the price of the item as the revenue. As a result, a marketing strategy has two elements: the sequence in which we offer the item to buyers, and the prices that we offer. In general it is advantageous to get influential buyers to buy the item early in the sequence; it even makes sense to offer such buyers smaller prices to get them to buy the item. We now describe our results:

Symmetric Settings. We start by studying a symmetric setting where all the buyers appear (ex-ante) identical to the seller, both in terms of the influence they exert and their response to offers.

In such a settings, the sequence in which to offer prices is immaterial and we can derive the optimal pricing policy using a *dynamic programming* (See Section 3.1). The optimal marketing strategy demonstrates the following behavior: the probability of buyers accepting their offer decreases as the marketing strategy progresses. Initially, the optimal marketing strategy offers discounts in an attempt to get buyers

to buy the item. This increases the value that buyers later in the sequence have for the item. This allows the optimal strategy to extract more revenue from subsequent buyers. In fact, early in the sequence the optimal strategy even gives away the item for free.

General Settings. Next, we consider algorithms to find the optimal marketing strategy in general settings. We first show that finding the optimal marketing strategy is NP-Hard by reduction from the maximum feedback arc set problem (See Section 3.2). This motivates us to consider approximation algorithms.¹

We identify a simple marketing strategy, called the *influence-and-exploit* strategy. Recall that any marketing strategy has two aspects: pricing and finding the right sequence of offers. In the initial *influence step*, motivated by the the form of the optimal strategy in the symmetric case, the seller starts by giving the item away for free to a specifically chosen set of players $A \subseteq V$. In the exploit step, the seller visits the remaining buyers ($V \setminus A$) in a random sequence and attempts to maximize the revenue that can be extracted from each buyer by offering it the (myopic) optimal price; note that this effectively ignores the influence that buyers in the set $V \setminus A$ exert on each other. (Note that the buyers in the set A , that we give the item away free to, are similar to *opinion leaders* [16] from the social contagion literature.)

We first show (See Section 4.1) that such strategies are a reasonable approximation of the optimal marketing strategy, which, by a hardness result is not polynomial-time computable. This is surprising because of the relative simplicity of influence-and-exploit strategies, which only uses two prices (the price zero and the optimal (myopic) price) and does not attempt to find the right offer sequence (it visits buyers in a random sequence).

This justifies studying the computational problem of finding the optimal influence-and-exploit strategy. In Section 4.2, show that if certain player specific revenue functions are submodular, then the expected revenue as a function of the set A is also submodular (Lemma 4.3). But as the revenue function is not monotone, we cannot use the simple greedy strategies suggested by Nemhauser, Wolsey and Fisher [13]. Instead, we use recent work by Feige, Mirrokni, and Vondrak [7] for maximizing *non-monotone* submodular functions, that gives a deterministic local search $\frac{1}{3}$ -approximation algorithm, and a randomized local search 0.4-approximation algorithm for this problem (Theorem 3).

1.2 Related work

Our work is inspired partly by the study of *Social Contagion* in the mathematical social sciences and, more recently, in computer science. Social contagion studies the dynamics of adoption of ideas or technologies in social networks. See Chapter 24 of [10] and the references therein. Typically, these works propose models for the process by which people in a social network adopt a new technology or idea. Kempe, Kleinberg, and Tardos [9] study the algorithmic question (posed by Domingos and Richardson [6]) of identifying a set of influential nodes in a social network: Assuming that the

¹An algorithm is a c -approximation if its revenue is at least c times the revenue of the optimal marketing strategy.

seller decides to give away k copies of an item, the question is to find a subset of k nodes in the network such that the subsequent adoption of the good is maximized; the value of k is externally specified.

As maximizing the spread of influence is often a means to an end rather than an end in itself, we consider marketing strategies that maximize revenue. While social contagion models are adequate for the study of the spread of a free good or service across society, they do not discuss the dependence of adoption on price, which makes studying revenue maximization hard in this setting. Our model defines the dependence of adoption on influence *and* price. Further, our model makes it possible to discuss *how many* people the item should be given away free to.

There has also been work by economists that studies the relationship of network externalities and pricing. These works are not algorithmically motivated. For instance, [18] studies the effect of network topology on a monopolist's profits from selling a networked good. Further, [5], studies a multi-round pricing game. As the rounds proceed, the seller may lower his price in an attempt to price discriminate and attract low value buyers. Their main result shows that early-round discounting motivated by network externalities can overwhelm the aforementioned tendency toward lower prices in later rounds and result in an ascending price over time.

Finally, as we pointed out earlier, in the influence maximization problem formalized in [9], the authors use the analysis of greedy algorithm for maximizing *monotone* submodular functions [13]. However, in our settings, the problem of optimal influence-and-exploit strategy is a *non-monotone* submodular function maximization; therefore, we make use of the recently developed local search algorithms for approximately maximizing non-monotone submodular functions.

2. PRELIMINARIES

In this section we discuss influence models, valid selling strategies and upper bounds on the maximum revenue that a seller can make.

Consider a seller who wants to sell a good to a set of potential buyers, V . The cost of manufacturing a unit of the good is zero and the seller has an unlimited supply of the good. We assume that the seller is a monopolist and is interested in maximizing its revenue. We start by discussing the well-known, optimal selling strategy in the (standard) setting with *no externalities*. As buyers do not influence each other, the seller can consider each buyer separately. We assume that though the seller does not know the buyer's exact value (maximum willingness to pay), it does know the distribution F from which its values are drawn; F is the cumulative distribution of the buyer's valuation, i.e., $F(t)$ the probability the buyer's value is less than t . We now define optimal pricing strategy (See for instance Myerson [12]).

Definition 1. Suppose that the player's value is distributed according to the distribution F . The optimal price p^* maximizes the expected revenue extracted from buyer i , i.e., the price p^* maximizes $p \cdot (1 - F(p))$. The optimal revenue is $p^* \cdot (1 - F(p^*))$ (in expectation).

2.1 Influence Model

We now describe a general setting where the buyer's influence each other; we also list concrete instances of this model. A buyer i 's value for the good now depends on the set of buyers that already own the good. It is determined by the function $v_i : 2^V \rightarrow R^+$; suppose this is a set $S \subseteq V \setminus \{i\}$, the value of buyer i is a non-negative number $v_i(S)$. When the social network is modeled by a graph, $v_i(\cdot)$ is a function only of neighbors of i in the graph.

Again, as in the setting with no externality, we assume the buyer knows the distributions from which the values are drawn; we treat the quantities $v_i(\cdot)$ as random variables. The seller knows the distributions of $F_{i,S}$ of the random variables $v_i(S)$, for all $S \subseteq V$ and for all $i \in V$. We assume throughout the paper that buyers' values are distributed independently of each other. Here are some concrete instantiations of this model that we study in the paper:

Uniform Additive Model In the uniform additive model, there weights w_{ij} for all $i, j \in V$. The value $v_i(S)$, for all $i \in V$ and $S \subseteq V \setminus \{i\}$, is drawn from the uniform distribution $[0, \sum_{j \in S \cup \{i\}} w_{ij}]$.

Symmetric Model In the symmetric model, the valuation $v_i(S)$ is distributed according to a distribution F_k , where $k = |S|$. (Note that the identities of the buyer i and the set S do not play a role.)

Concave Graph Model In this model, each buyer $i \in V$ is associated with a non-negative, monotone, concave function $f_i : R^+ \rightarrow R^+$. The value $v_i(S)$ for all $i \in V$, $S \subseteq V \setminus \{i\}$, is equal to $f_i(\sum_{j \in S \cup \{i\}} w_{ij})$. Each weight w_{ij} is drawn independently from a distribution F_{ij} . The distributions $F_{i,S}$ can be derived from the distributions F_{ij} for all $j \in S$.

We will discuss these models along with other possible models in details in Section 5.

2.2 Marketing Strategies

As discussed in the introduction, when buyers influence each other, the seller can conduct sales in an intelligent sequence and offer intelligent discounts so as to optimize its revenue. In this section we formally describe the space of possible selling strategies.

A *marketing strategy* has the seller visiting buyers in some sequence and offering each buyer a price. Each buyer either accepts (buys the item and pays the offered price) or rejects (does not buy and does not pay the seller) the item; we assume that each buyer is considered exactly once. Both the prices offered and the sequence in which buyers are visited can be adaptive, i.e., they can be based on the history of accepts and rejects. A marketing strategy thus identifies the next buyer to visit and the price to offer it as a function of the history. Throughout this paper, buyers are assumed to be *myopic*, i.e., they are influenced only by buyers who have already bought the item. At any point in time, if a set S of buyers already owns the item, the value of buyer i is $v_i(S)$.

A *run* of a marketing strategy consists of sequence of offers, one to each buyer in V along with the set of accepted and rejected offers. The revenue from the run is the sum of

the payments from the accepted offers. A marketing strategy and the value distributions together yield a distribution over runs—this defines the expected revenue of the marketing strategy. We call the marketing strategy that optimizes revenue the *optimal marketing strategy*.

2.3 An Upper Bound on Revenue

In this section we discuss why using the optimal price (Definition 1) is short-sighted. We also derive an upper bound on the revenue of the optimal marketing strategy.

Suppose that the seller visits a specific buyer i at some point in a run and a set S of buyers has already bought the good. The value of the buyer i is now distributed as $F_{i,S}$. What price should the seller offer to the buyer? We note that optimal pricing (Definition 1) is no longer optimal; we may want to offer the buyer a discount, so that it buys the item and influences others. However if the seller is myopic and ignores the buyer i 's ability to influence other buyers it would offer the optimal price; motivated by this, we henceforth refer to the optimal price as the *optimal (myopic) price*.

We finish the section by deriving an upper bound on the revenue of the optimal marketing strategy in terms of certain player specific revenue functions. Let $R_i(S)$ be the revenue one can extract from player i given that set S of players have bought the item using the optimal (myopic) price (See Definition 1). Naturally, R_i is *non-negative*. We assume that the functions R_i are monotone, i.e. for all i and $A \subseteq B \subseteq V \setminus i$, $R_i(A) \leq R_i(B)$; this implies that buyers only exert positive influence on each other. Monotonicity of the revenue functions implies the following upper bound on the revenue of the optimal marketing strategy.

Fact 1. The revenue of the optimal marketing strategy is at most is at most $\sum_{i \in V} R_i(V)$.

In Section 4, we additionally assume that R_i is *submodular* (for all i , for all $A \subset V$ and $B \subset V \setminus \{i\}$, $R_i(A \cup B) + R_i(A \cap B) \leq R_i(A) + R_i(B)$). Submodularity is the set analog of concavity: it implies that the marginal influence of one buyer on another decreases as the set of buyers who own the good increases. We further discuss the submodularity assumption in Section 5.

2.4 Additional Technical Preliminaries

In this section we list some technical facts that we use in the paper. We repeatedly use the following fact about monotone submodular functions. We leave its proof to the appendix.

LEMMA 2.1. *Consider a monotone submodular function $f : 2^V \rightarrow \mathbb{R}$ and subset $S \subset V$. Consider random set S' by choosing each element of S independently with probability at least p . Then $E[f(S')] \geq p \cdot f(S)$.*

Some of our results rely on the value distributions satisfying a certain *monotone hazard rate* condition. We first define the hazard rate function of a distribution.

Definition 2. The hazard rate h of a distribution with a density function f , distribution function F and support $[a, b]$ is $h(t) = \frac{f(t)}{1-F(t)}$. The distribution function can be expressed in terms of the hazard rate: $F(t) = 1 - e^{-\int_a^t h(x)dx}$.

Definition 3. A distribution, with a density function f and distribution function F , satisfies the monotone hazard rate condition if and only if for any point t in the support, $h(t) = \frac{f(t)}{1-F(t)}$ is monotone non-decreasing.

The assumption that the values distribution satisfies the monotone hazard rate condition is a fairly weak. Such an assumption is commonly employed in auction theory [12] to model value distributions—several distributions such as the uniform, the exponential, the normal distribution satisfy this condition. For instance, the uniform distribution in the interval $[0, 1]$ has a hazard rate $\frac{1}{1-t}$. We further discuss the monotone hazard rate assumption in Section 5.

3. OPTIMAL MARKETING STRATEGIES

3.1 Symmetric Settings

In this section, we study symmetric settings, and show that we can identify the optimal marketing strategy based on a simple dynamic programming approach. We assume that buyer values are defined according to the symmetric model from the previous section, where the buyer values are drawn from one of $|V|$ distributions F_k .

We now derive the optimal marketing strategy. As the model is completely symmetric in the buyers, the sequence in which it visits buyers is irrelevant. Further, the offered prices are a function only of the number of buyers that have accepted and the number of buyers who have not, as yet, been considered. Let $p(k, t)$ be the offer price to the buyer under consideration, used by the optimal marketing strategy, given that k people have bought the good and t buyers are not as yet considered (including the buyer currently under consideration); and $R(k, t)$ is the maximum expected revenue that can be collected from these remaining buyers. We now set-up and solve a recurrence in terms of the variables p and R . We assume that the density function of the distribution F_k , $f_k(S)$, exists.

Given a price p , if the buyer accepts, we can collect the revenue of $p + R(k+1, t-1)$, and if it rejects, we can collect revenue of $R(k, t-1)$. Moreover, the buyer accepts if and only if its value is at least p , i.e with probability $1 - F_k(p)$.

As a result, we have to set the price p to maximize the expected remaining revenue. For any price p , the expected remaining revenue is:

$$F_k(p) \cdot R(k, t-1) + (1 - F_k(p)) \cdot (R(k+1, t-1) + p)$$

The optimal price can be found by differentiating the above expression with respect to p and setting to 0:

$$f_k(p)(R(k, t-1) - R(k+1, t-1) - p) + 1 - F_k(p) = 0$$

We can then set $p(k, t)$ to the value which satisfies the above equation. The variable $R(k, t)$ is now easy to compute. The above dynamic program can be solved in time quadratic in the number of buyers. For the base case, note that $R(k, 0) = 0$. This defines the optimal marketing strategy; note that all we need is for the density functions to exist, there were no additional assumptions in the analysis. We now state the main result of this section (without proof):

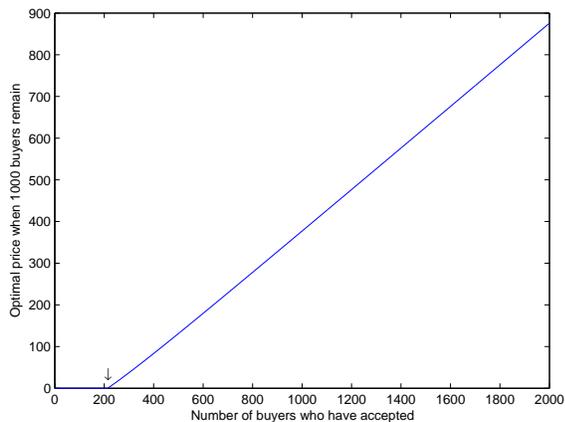


Figure 1: The optimal price for additive influence function when 1000 buyers remain changing the number of buyers who have accepted. The arrow shows the place at which the optimal price becomes nonzero.

LEMMA 3.1. *In the symmetric influence model, the optimal strategy can be computed in polynomial time.*

We conclude the section by briefly investigating a concrete symmetric setting: Suppose the value of agent i with S served, $v_i(S)$, is uniform $[0, |S| + 1]$. (A symmetric setting where the distribution F_k is the uniform distribution on $[0, k + 1]$.) Figures 1 and 2 depict the variation in the optimal price as k and t vary; Figure 1 confirms that for a fixed t , the optimal price increases as the number of buyers who have already bought the item increases. Figure 2 confirms that for a fixed k , as the number of players who remain goes up, it makes more sense to ensure that the player under consideration buys the good even if this means sacrificing the revenue earned from the player. Both monotonicity properties hold more generally. Figure 2 also shows that at the beginning of the marketing strategy, when a large number of buyers remain in the market, the optimal price is zero. This observation motivates studying the *influence-and-exploit* marketing strategy discussed in Section 4.

3.2 Hardness

We now consider the algorithmic problem of finding optimal marketing strategies in general settings. In this section, we show that the problem of computing the optimal strategy is NP-Hard even when there is *no uncertainty* in the input parameters. In particular, we assume that the values $v_i(S)$ are precisely known to the seller; all the distributions $F_{i,S}$ are degenerate point distributions. In such a setting it is easy to see that the only problem is to find the right sequence of offers. Given any offer sequence, the prices to offer are obvious; if a set S of buyers have previously bought, offer the next buyer i price $v_i(S)$. This price simultaneously extracts the maximum revenue possible and ensures that the buyer buys and hence exerts influence on future buyers. We now show that finding the optimal sequence is NP-Hard even when the values are specified by a simple additive model. We consider the *additive* model where, $v_i(S) = \sum_{j \in S \cup \{i\}} w_{ji}$.

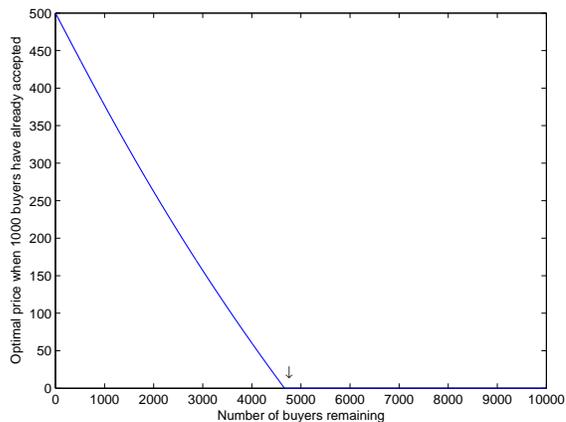


Figure 2: The optimal price for additive influence function when 1000 buyers have accepted changing the number of remaining buyers. The arrow shows the place at which the optimal price becomes zero.

The reduction is from the *maximum feedback arc-set* problem; the proof is in the appendix.

LEMMA 3.2. *Finding the optimal marketing strategy is NP-hard even with complete information about buyer values.*

The above hardness result shows that even with full information about the players' values, computing the optimal ordering is hard. Motivated by this hardness result, we design *approximately* optimal marketing strategies that can be found in polynomial time. As the above reduction is approximation preserving, to achieve better than $1/2$ -approximation for our problem, we must improve the approximation factor of the maximum feedback arc set problem. The best approximation algorithm known for the maximum feedback arc set problem is a $\frac{1}{2}$ -approximation algorithm [4, 8], and it is long-standing open question to achieve better than $\frac{1}{2}$ -approximation for. As our problem also involves the pricing aspect, we shall content ourselves with trying to get close to the benchmark of $1/2$. In the appendix we include an example that demonstrates the importance of computing the right offer sequence even in an undirected setting.

4. INFLUENCE-AND-EXPLOIT MARKETING

Motivated by the hardness result from Section 3.2, we now turn our attention to designing polynomial-time algorithms that find approximately optimal marketing strategies. Recall that a marketing strategy broadly has two elements, the offer sequence and the pricing. We identify a simple, effective marketing strategy, called the *influence-and-exploit (IE) strategy*. We start by motivating this strategy, then show that it is effective in a very general sense and finish by discussing techniques to find optimal strategies of this form.

We now motivate the structure of the IE strategy; the strategy has an *influence step*, which gives the item away for free to a judiciously selected set of buyers; followed by

an *exploit step* that is based on a random sequence of offers and a greedy pricing strategy.

1. The optimal marketing strategy in the symmetric setting started by giving the item away for free to a significant fraction of the players; this motivates the influence step.
2. In the previous section we noted that the best known approximation algorithm for the maximum feedback arc-set problem is a 1/2-approximation; surprisingly, picking a random sequence of nodes yields this (As each edge is selected with probability 1/2). Inspired by this, during the exploit step, we will visit buyers in a sequence picked uniformly at random.
3. We will use optimal (myopic) pricing (See Definition 1) in the exploit step; we will attempt to maximize revenue extracted from a buyer, without worrying about the influence that it exerts on others.

We now define the **IE Strategy**. The strategy has two steps:

1. **Influence:** Give the item free to buyers in a set A .
2. **Exploit:** Visit the buyers of $V \setminus A$ in a sequence σ (picked uniformly at random from the set of all possible sequences). Suppose that a set $S \subseteq V \setminus \{i\}$ of buyers have already bought the item before buyer i is made an offer. Offer buyer i the optimal (myopic) price as a function of the distribution $F_{i,S}$. Note that the optimal (myopic) price is adaptive, and is based on the history of sales.

Though, we do not extract any revenue from set A , we are guaranteed that these buyers accept the item and influence other buyers. This will allow us to extract added revenue from the set $V \setminus A$ of buyers that more than compensates for the initial loss in revenue. There are two issues: How good is the IE strategy compared to the optimal strategy? What set A maximizes revenue? The next two sections answer these questions.

4.1 How Good are Influence-and-Exploit Strategies?

Note that IE strategies are fairly simple (they only use two extreme prices and random orderings) and it is not clear how much we lose, restricting our attention to this class of strategies. In this section we show that they compare favorably to the optimal revenue-maximizing strategy. Before stating improved approximation guarantees for various settings, we observe the following simple fact:

Remark 1. Given any set of submodular revenue functions R_i , the expected revenue from the optimal IE strategy is at least $\frac{1}{4}$ of the optimal revenue.

PROOF. We can prove this remark by taking the set A of the IE strategies to be a random subset of buyers where each buyer is chosen independently with probability $\frac{1}{2}$. By Lemma 2.1, the expected revenue from this IE strategy is at least $\sum_{i \in V \setminus A} R_i(A) = \sum_{i \in V \setminus A} \frac{R_i(V)}{2}$. Since each buyer

is in set $V \setminus A$ with probability $\frac{1}{2}$, the expected revenue of this strategy is at least $\sum_{i \in V} \frac{R_i(V)}{4}$. By Fact 1, the expected revenue of this IE strategy is a $\frac{1}{4}$ -approximation of the optimal revenue. \square

Now, we prove several improved approximation guarantees for IE strategies for special classes of the problem. For the concrete setting studied at the end of Section 3.1, it is possible to show that the best IE strategy is a 0.94-approximation to the optimal revenue. We now analyze the IE strategy in the undirected additive model (See Section 2.1). We show that there exists an IE strategy that gives a $\frac{2}{3}$ -approximation algorithm for this problem. We start by stating an easy fact about such uniform distributions:

Fact 2. Suppose a buyer has value distributed uniformly in an interval $[0, M]$, then the optimal (myopic) price is $M/2$, which is also the mean of the distribution. The optimal (myopic) revenue is $M/4$.

We now describe the IE strategy. All we need to specify is the set A . Let $\frac{N = \sum_{i \in V} w_{ii}}{2}$ and $\frac{E = \sum_{\{ij\}, i \neq j} w_{ij}}{2}$. Let $q = \frac{E - 2N}{3E}$. Let A be a random subset of nodes where each node is sampled with probability q .

THEOREM 1. *In the undirected, additive model, IE with the set A constructed as above yields at least $\frac{2}{3}$ of the maximum possible revenue.*

PROOF. We start by showing an upper-bound on the revenue that any strategy can attain. The upper bound is tighter than the bound from Fact 1; we use the observation that only one of w_{ij} or w_{ji} for $i \neq j$, can contribute to the revenue. For any strategy, fix the order in which the sales happened. Even assuming that *every* buyer buys the item, by Fact 2 the revenue extracted from the i th bidder in the sequence is $1/4 \cdot \left(w_{ii} + \sum_{j \in S_{i-1}} w_{ji} \right)$; here S_k is the first k players in the ordering. Summing over the bidders we have that the optimal revenue is at most $1/2 \cdot (N + E/2)$. Let T_i be the set of buyers who buy the item before buyer v . T_i includes A , and a random subset of $V \setminus A$. Thus, for any buyer v , a buyer u is in set T_i with probability $q + \frac{(1-q)}{4} = \frac{1+3q}{4}$. Thus, for any buyer $i \in V \setminus A$, $E[v_i(T_i)] = w_{ii}/2 + \sum_{j \neq i} \frac{1+3q}{8} w_{ji}$, thus the expected revenue from $i \in V \setminus A$ is $\frac{1}{2} E[v_i(T_i)] = \frac{1}{4} w_{ii} + \sum_{j \neq i} \frac{1+3q}{16} w_{ji}$. Moreover, a buyer v in set $V \setminus A$ with probability $1 - q$. As a result, the expected revenue of the above algorithm is at least

$$\begin{aligned} \frac{1}{2} \sum_{i \in V} (1-q) E[v_i(T_i)] &= \sum_{i \in V} (1-q) \left(\frac{w_{ii}}{4} + \sum_{j \neq i} \frac{1+3q}{16} w_{ji} \right) \\ &= \frac{1}{4} \sum_{i \in V} (1-q) w_{ii} + \sum_{\{i,j\}, j \neq i} \left(\frac{1+2q-3q^2}{16} \right) w_{ji} \end{aligned}$$

Thus, the expected revenue is at least $\frac{1}{2}(1-q)N + \left(\frac{1+2q-3q^2}{8} \right) E$. In order to maximize the expected revenue, we should set: $q = \frac{E-2N}{3E}$. For this value of q , the expected revenue is at least $\frac{(E+N)^2}{6E} \geq \frac{(E^2+2EN)}{6E} \geq \frac{E}{6} + \frac{N}{3}$. This proves the theorem. \square

We now show that IE strategies compare favorably to the optimal strategy even in a fairly general setting—the revenue functions are submodular, monotone and non-negative and the value distributions satisfy the monotone hazard rate condition. We start by showing that if the value distribution satisfies the monotone hazard rate condition, the buyer accepts the optimal (myopic) price with a constant probability.

LEMMA 4.1. *If value distribution satisfies the monotone hazard rate condition, the buyer accepts the optimal (myopic) price with probability at least $1/e$.*

PROOF. By Definition 2, $1 - F(t) = e^{-\int_a^t h(x)dx}$. As F_i satisfies the monotone hazard rate condition, $1 - F(t) \geq e^{-\int_a^t h(t)dx}$. At the optimal price, we have that $1/t = h(t)$. So $1 - F(x) \geq e^{-\int_a^x 1/t dx} = e^{-\frac{t-a}{t}} \geq 1/e$, as e^x is a monotone function. \square

We now use the above lemma to prove the following theorem.

THEOREM 2. *Suppose that the revenue functions $R_i(S)$, for all $i \in V$ and $S \subseteq V \setminus \{i\}$ are monotone non-negative and submodular and the distributions $F_{i,S}$ for all $i \in V$ and $S \subseteq V \setminus \{i\}$ satisfy the monotone hazard rate condition. Then there exists a set A for which the IE strategy is a $\frac{e}{4e-2}$ -approximation of the optimal marketing strategy.*

PROOF. Let A be a random subset of buyers where each buyer is picked with probability p . Consider the IE strategy for this set A . For a buyer $i \in V \setminus A$, let T_i be the random subset of buyers who have bought the item before buyer i . Each buyer j is in $V \setminus A$ with probability $1 - p$, it appears before i with probability $\frac{1}{2}$, and in this case, j buys the item by probability at least $\frac{1}{e}$ (from Lemma 4.1), thus, each buyer $j \in V \setminus A$ is in set T_i with probability at least $\frac{1-p}{2e}$. Also each buyer j is in A with probability p in which $j \in T_i$ as well. As a result, each buyer $j \in V$ is in T_i with probability at least $p + \frac{1-p}{2e}$.

Let R_i be the expected revenue from buyer i in this algorithm. Then by monotonicity and submodularity of the expected revenue function R_i , and by Lemma 2.1, the expected revenue from T_i is at least $(p + \frac{1-p}{2e})R_i(V)$. Thus, the expected revenue from this algorithm is at least $(p + \frac{1-p}{2e})\sum_{i \in V \setminus A} R_i(V)$. Since each buyer i is in $V \setminus A$ with probability $1 - p$, the expected revenue from the IE strategy is at least $(1 - p)(p + \frac{1-p}{2e})\sum_{i \in V} R_i(V)$ which is maximized by setting $p = \frac{e-1}{2e-1}$. The theorem follows from Fact 1. \square

4.2 Finding Influence-and-Exploit Strategies

In the previous section, we showed that in various settings influence and exploit strategies approximate the optimal revenue within a reasonable constant factor. Motivated by this, we attempt to find good IE strategies in more general settings. What set A of buyers, should we initially give the item for free so that the revenue from the subsequent exploit stage is maximized? In other words, we want to find a set A that maximizes $g(A)$ where $g(A)$ is the expected revenue of the IE strategy when we give the item for free to set A in the first step. Though we do not compute optimal set

A , we compute an A that gives a good approximation. The main result of this section is the following:

THEOREM 3. *There is a deterministic polynomial-time algorithm that computes a set A , such that the revenue of the IE strategy with this set yields at least a $\frac{1}{3}$ -fraction of the revenue of the optimal IE strategy. Moreover, there exists a randomized polynomial-time 0.4-approximation algorithm for the optimal IE strategy.*

We now describe the deterministic algorithm mentioned in the above theorem. It is based on a local search approach.

Local Search

1. Initialize set $A = \{v\}$ for the singleton set $\{v\}$ with the maximum value $g(\{v\})$ among singletons.
2. If neither of the following two steps apply (there is no local improvement), output A .
3. For any buyer $i \in V \setminus A$, if $g(A \cup \{i\}) > (1 + \frac{\epsilon}{n^2})g(A)$ (adding an element to A increases revenue), then set $A := A \cup \{i\}$ and go to 2.
4. For any buyer $i \in A$, if $g(A \setminus \{i\}) > (1 + \frac{\epsilon}{n^2})g(A)$ (deleting an element from A increases revenue), then set $A := A \setminus \{i\}$ and go to 2.

Since at each step of the local search algorithm, the expected revenue improves by a factor of $(1 + \frac{\epsilon}{n^2})$, and the initial value of $g(A)$ is at least $\frac{1}{n}$ of the maximum value, the number of local improvements of this algorithm is at most $\log_{(1 + \frac{\epsilon}{n^2})} 4 = O(\frac{n^2}{\epsilon})$; this is also an explanation for why the algorithm necessarily terminates. Further, we can compute $g(A)$ for any set A in polynomial time by sampling a polynomial number of scenarios, and taking the average of the function for these samples. This shows that the above algorithm runs in polynomial time.

The proof of Theorem 3 follows from the following more general result by Feige, Mirrokni, and Vondrak [7] about the use of the local search algorithm (above) in maximizing non-monotone submodular functions. Though we omit the details, there is a more complicated randomized algorithm that can be used in place of the deterministic local search algorithm to get a slightly better approximation ratio [7].

LEMMA 4.2. [7] *Suppose the set function $g(\cdot)$ is non-negative and submodular. Let M be the maximum value of the submodular set function. Then the deterministic local search algorithm finds a set A such that $g(A) \geq \frac{1}{3}M$. Moreover, there exists a randomized local search algorithm that finds a set A such that $g(A) \geq \frac{2}{5}M$.*

Given the above theorem, to complete the proof of Theorem 3, it is sufficient to show that the function $g(A)$ is non-negative and submodular. In order to prove submodularity of function g , we use the following facts about submodular functions.

Fact 3. If f and g are submodular, for any two real numbers α and β , the set function $h : 2^V \rightarrow R$ where $h(S) = \alpha f(S) + \beta g(S)$ is also submodular. The set function h where $h(S) = f(V \setminus S)$ is submodular. For a fixed subset $T \subset V$, function h where $h(S) = f(S \cup T)$ is also submodular.

We now show that under certain conditions on the revenue functions R_i for $i \in V$, the set function $g(A)$ is a non-negative submodular function.

LEMMA 4.3. *If all the revenue functions R_i for $i \in V$ are non-negative, monotone and submodular, then the expected revenue function $g(A) = \sum_{i \in V \setminus A} R_i(A)$ is a non-negative submodular set function.*

PROOF. It is easy to see that g is non-negative for all i . We focus on proving that g is submodular: We need to prove that for any set $A \subseteq V$ and $C \subseteq V$:

$$g(A) + g(C) \geq g(A \cup C) + g(A \cap C),$$

First, using monotonicity of R_i , for each $i \in (A \setminus C) \cup (C \setminus A)$:

$$\sum_{i \in A \setminus C} R_i(C) + \sum_{i \in C \setminus A} R_i(A) \geq \sum_{i \in A \setminus C} R_i(A \cap C) + \sum_{i \in C \setminus A} R_i(A \cap C) \quad (1)$$

Now, using submodularity of R_i , for each $i \in V \setminus (A \cup C)$,

$$R_i(A) + R_i(C) \geq R_i(A \cup C) + R_i(A \cap C).$$

Therefore, summing the above inequality for all $i \in V \setminus (A \cup C)$, we get:

$$\begin{aligned} & \sum_{i \in V \setminus (A \cup C)} R_i(A) + \sum_{i \in V \setminus (A \cup C)} R_i(C) \\ & \geq \sum_{i \in V \setminus (A \cup C)} R_i(A \cup C) + \sum_{i \in V \setminus (A \cup C)} R_i(A \cap C) \end{aligned}$$

Summing equations 1, 2,

$$\begin{aligned} & \sum_{i \in V \setminus A} R_i(A) + \sum_{i \in V \setminus C} R_i(C) \geq \\ & \sum_{i \in V \setminus (A \cup C)} R_i(A \cup C) + \sum_{i \in V \setminus (A \cap C)} R_i(A \cap C), \end{aligned}$$

This proves the result. \square

Note that function g is not monotone and so we cannot use the simple greedy algorithm developed by Nemhauser, Wolsey, and Fischer [13], also used by Kempe, Kleinberg, Tardos [9]. Instead, we need to use the local search and randomized algorithms developed by Feige, Mirrokni, and Vondrak [7].

5. DISCUSSING OF THE MODEL

In this section, we discuss the validity of the modeling assumptions made in Section 4. We do so by discussing the *concave graph model* from Section 2. After justifying the concave graph model, we show that it satisfies the submodularity and the monotone hazard assumptions from the previous section.

Recall that in this model where the uncertainty is in the influence that a buyer has on another buyer and the influences are combined using buyer specific concave functions. The concavity models the diminishing returns that one expects the influence function to have. Such concavity has also been demonstrated by empirical studies: [1] studies the effect of influence on joining an online community; what is

the probability of joining an online community given that n of your friends were already members. They show that the probability increases almost logarithmically. (See Figure 24.1 in [10]). Such concave influence functions have another implication: once sufficiently many buyers have bought the item, it is easy to see that additional sales have little influence. From this point on it is optimal to use optimal (myopic) prices. In particular, if buyers are relatively symmetric, optimal (myopic) pricing can be implemented via a posted price.

It may be possible to use the link structure of online social networks to estimate w_{ij} . Studies such as [1] can be used to determine the precise form of the functions f_i . In practice, we could reduce the parameters that need to learn by making intelligent symmetry assumptions. For instance, it might be reasonable to assume that there are two categories of buyers, buyers who wield considerable influence (opinion leaders) and other buyers.

We now discuss the validity of the assumptions made about the player specific revenue functions, namely non-negativity, monotonicity and submodularity. Non-negativity is obvious. Monotonicity follows from the non-negativity of the weights and the non-negativity and monotonicity of f_i . We now show that the means of the values, $v_i(\cdot)$, are submodular.

LEMMA 5.1. *In the concave graph model, the expected value of the random variable $v_i(S)$, $\bar{v}_i(S)$ is a monotone, non-negative, submodular set function.*

PROOF. Fix a buyer i . Condition on the values of the random variables w_{ij} . For any subsets $S \subseteq S' \subseteq V$ and buyer k not in S' , we claim that:

$$(v_i(S \cup \{k\}) - v_i(S)) - (v_i(S' \cup \{k\}) - v_i(S')) \geq 0$$

This follows from the concavity of f_i . Thus the function $v_i(\cdot)$ is point-wise submodular. We can now use Fact 3 to complete the proof. \square

Though we cannot quite prove that the player-specific revenue functions are submodular (essentially revenue does not allow for a simple point-wise argument as above), we conjecture that this is true; it is easy to prove the conjecture in a setting where, for a fixed buyer i , the random variables $v_i(S)$ for all $S \subseteq V \setminus \{i\}$ are identically distributed up to a scale factor; note that this is a generalization of the additive model from Section 2.1.

We now argue why it is reasonable to assume that the value distributions satisfy the monotone hazard rate condition. First in many situations, we may expect a significant fraction of the value of a buyer i to be independent of external influence (w_{ii} dominates w_{ij} for $i \neq j$); in such cases the monotone hazard rate assumption is commonly made in auction theory. Second, by the well-known Central Limit Theorem, the sum of the independently distributed influence variables (w_{ij} s for some fixed i) will be approximately like a normal distribution, so long as the variables are roughly identically distributed; it is known that the normal distribution satisfies the monotone hazard rate condition. Finally, we can use the following closure properties of the monotone hazard rate condition to show that if the distributions F_{ij}

satisfy the monotone hazard condition, then so do the value distributions $F_{i,S}$.

LEMMA 5.2. *Fix an arbitrary buyer $i \in V$. In the concave graph model, if the distributions F_{ij} satisfy the monotone hazard rate condition for all j , then for all sets $S \subseteq V$, the distributions $F_{i,S}$ satisfies the monotone hazard rate condition.*

We use the following lemma established in [2]. The lemma (proof omitted) formalizes the fact that the distribution of the sum of the random variables is only better concentrated than the distributions of the individual variables.

LEMMA 5.3. [2] *The monotone hazard rate condition is closed under addition in the following sense: For any set of random variables a_j , if each a_j is drawn from a distribution that satisfies the monotone-hazard-rate condition, then the random variable $\sum_j a_j$ also satisfies the monotone hazard rate condition*

The next lemma (proof in the appendix) shows that the monotone hazard rate condition is closed under the application of a monotone function.

LEMMA 5.4. *If a random variable a is drawn from a distribution (with cumulative distribution function F and density function f) that satisfies the monotone hazard rate condition, then the random variable $h(a)$ (with distribution F_h and a density function f_h) also satisfies the monotone hazard rate condition, so long as h is strictly increasing.*

We now finish the proof of Lemma 5.2. By Lemma 5.3, the random variable $\sum_{i \in S \cup \{i\}} w_{ij}$, satisfies the monotone hazard rate condition. By Lemma 5.4, and as f_i is increasing, we have the proof.

Finally, though we assume throughout the paper that optimal myopic prices can be calculated, we note that it is also reasonable to use mean values instead; the IE strategy thus modified will continue to give a constant factor approximation, though the constant is somewhat worse. The key lemma (Lemma A.1) which makes this possible is stated in the appendix; this lemma plays the role of Lemma 4.1.

6. CONCLUSION

In this paper, we discuss the optimal pricing strategies in social networks considering that the valuation of the digital good for users depends on other users using a service. We considered the incomplete information setting in which we only need to know the optimal (myopic) price. Our main contribution in the paper is identifying a family of IE strategies, proving that they provide improved approximation algorithms, and finally, computing a good IE strategy. Some open questions:

1. In Section 4.2, we discuss a local search algorithm that yields a 0.33 approximation for computing the optimal influence set. There is a more involved randomized algorithm [7] that yields a 0.4-approximation algorithm. It has been shown that no polynomial-time algorithm can achieve an approximation factor of 0.5 for maximizing general monotone-submodular functions? Are

there better algorithms possible for the special cases considered in this paper?

2. Are there other strategies that can be computed in polynomial time that yield better revenue? For instance, can we use intelligently constructed sequences rather than random orderings?
3. It would also be interesting to develop pricing algorithms for a model where the seller does not visit buyers in a sequence, but simply posts prices; we expect that IE type strategies will continue to be effective in such settings. Finally, disallowing price discrimination and designing fixed-price mechanisms is also an interesting research direction.

7. REFERENCES

- [1] Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. Group formation in large social networks: membership, growth, and evolution. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 44–54, New York, NY, USA, 2006. ACM.
- [2] Barlow, Richard E. and Marshall, Albert W. Bounds for distributions with monotone hazard rate, i. *The Annals of Mathematical Statistics*, 35(3):1234–1257, sep 1964.
- [3] Barlow, Richard E. and Marshall, Albert W. Bounds for distributions with monotone hazard rate, ii. *The Annals of Mathematical Statistics*, 35(3):1258–1274, sep 1964.
- [4] Bonnie Berger and Peter W. Shor. Approximation algorithms for the maximum acyclic subgraph problem. In *SODA '90: Proceedings of the first annual ACM-SIAM symposium on Discrete algorithms*, pages 236–243, Philadelphia, PA, USA, 1990. Society for Industrial and Applied Mathematics.
- [5] Luis Cabral, David Salant, and Glenn Woroch. Monopoly pricing with network externalities. Industrial Organization 9411003, EconWPA, November 1994.
- [6] Pedro Domingos and Matt Richardson. Mining the network value of customers. In *KDD '01: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 57–66, New York, NY, USA, 2001. ACM.
- [7] Uriel Feige, Vahab S. Mirrokni, and Jan Vondrak. Maximizing non-monotone submodular functions. In *FOCS '07: Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 461–471, Washington, DC, USA, 2007. IEEE Computer Society.
- [8] Refael Hassin and Shlomi Rubinstein. Approximations for the maximum acyclic subgraph problem. *Information Processing Letters*, 51(3):133–140, 1994.
- [9] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge*

discovery and data mining, pages 137–146, New York, NY, USA, 2003. ACM.

- [10] J. Kleinberg. *Cascading behavior in networks: algorithmic and economic issues*. Cambridge University Press, 2007.
- [11] Jay Meattle. <http://blog.compete.com/2007/01/25/top-20-websites-ranked-by-time-spent/>.
- [12] R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.
- [13] G. Nemhauser, L. Wolsey, and M. Fisher. An analysis of the approximations for maximizing submodular set functions. *Mathematical Programming*, 14:265–294, 1978.
- [14] Alantha Newman. The maximum acyclic subgraph problem and degree-3 graphs. In *APPROX '01/RANDOM '01: Proceedings of the 4th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems and 5th International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 147–158, London, UK, 2001. Springer-Verlag.
- [15] Ed Oswald. http://www.betanews.com/article/Google_Buy_MySpace_Ads_for_900m/1155050350.
- [16] Everett Rogers. *Diffusion of Innovations, 5th Edition*. Free Press, August 2003.
- [17] Katharine Q. Seelye.
- [18] Pekka SÄD'Äd'skilähti. Monopoly pricing of social goods. MPRA Paper 3526, University Library of Munich, Germany, 2007.
- [19] Rob Walker. <http://www.slate.com/id/1006264/>.
- [20] Tim Weber. <http://news.bbc.co.uk/1/hi/business/6305957.stm?lsf>.

APPENDIX

A. PROOFS

Proof of Lemma 2.1.

PROOF. Fix an ordering σ of the elements of the set S . We can write $f(S)$ as the sum $\sum_{1 \leq i \leq |S|} f(S_i) - f(S_{i-1})$. Here S_i consists of the first i elements of the set S and we assume that $f(S_0) = 0$.

Recall the definition of the set S' from the lemma statement. Using linearity of expectations, we have that:

$$\begin{aligned} E[f(S')] &= E\left[\sum_{1 \leq i \leq |S'|} f(S'_i) - f(S'_{i-1})\right] \\ &\geq \sum_{1 \leq i \leq |S|} p \cdot (f(S_i) - f(S_{i-1})) \\ &= p \cdot f(S) \end{aligned}$$

The second inequality uses the submodularity of f . \square

Proof of Lemma 3.2.

PROOF. We show how to reduce any instance of the NP-Hard maximum feedback arc set problem [4, 8, 14] to our problem. This establishes that our problem is also NP-Hard and we cannot achieve a polynomial time solution to our problem unless $P = NP$.

In an instance of the maximum feedback arc set problem, given an edge-weighted directed graph, we need to order the nodes of the graph to maximize the total weight of edges going in the backward direction in the ordering. We now describe the reduction.

Let the nodes of the graph be the set of buyers. The edge weights are the weights w_{ij} . Let w_{ij} equal 0 for edges absent. We now define the pricing. Given the ordering in which to offer buyers, we offer prices equal to the player's value; for a player i it is $\sum_{j \in S \cup \{i\}} w_{ji}$, where S is the set of nodes visited before i . Given any ordering σ , the revenue from such pricing is equal to the weight of the feedback arc set when the nodes in the graph are ordered in the reverse of σ . Thus finding the the optimal marketing strategy is equivalent to computing the maximum feedback arc-set. \square

The above proof shows the importance of constructing the right offer sequence; we now observe that even in settings in which the influence is bidirectional, but the buyer has incomplete information, the offer sequence matters. For example, consider the additive model corresponding to a star graph of n buyers. Suppose that w_{ii} is 0, w_{ij} , $j \neq i$ is 0 if neither i or j is the center; and w_{ij} is drawn from the uniform distribution on the interval $[0, 1]$ otherwise. We find that the optimal marketing strategy starts at the center and offers it a carefully calculated price; then it offers the remaining buyers the optimal (myopic) price. Somewhat suprisingly, if, instead, we had complete information, the offer sequence does not matter. The example shows that incomplete information makes the offer sequence important.

LEMMA A.1. [3] *A buyer, whose value is distributed according to a distribution that satisfies the monotone hazard rate condition, accepts an offer price equal to the mean value with probability at least $1/e$.*

PROOF. Fix the set S of buyers who already own the item and the buyer under consideration, i . Let f and F be the density and distribution functions for the buyer's value $v_i(S)$. By Definition 2, we can write $\log(1 - F(x)) = -\int_a^x h(t)dt$. As $h(t)$ is non-decreasing in t , $\log(1 - F(x))$ is concave. Now, using Jensen's inequality, $\log(1 - F(\mu)) \geq \int_0^\infty \log(1 - F(x))dF(x) = \int_0^1 \log(1 - y)dy \geq -1$. (Replacing $F(x)$ by y .) Taking the exponent on both sides completes the proof. \square

Proof of Lemma 5.4.

PROOF. Because the function h is strictly increasing, the inverse function h^{-1} is defined. So for all t ,

$$\frac{f_h(t)}{1 - F_h(t)} = \frac{f(h^{-1}(t))}{1 - F(h^{-1}(t))}$$

Thus the monotone hazard rate condition is satisfied for the random variable $h(a)$ if and only if for all t and $e > 0$, $\frac{f(h^{-1}(t))}{1 - F(h^{-1}(t))} \leq \frac{f(h^{-1}(t+e))}{1 - F(h^{-1}(t+e))}$. But this is true as the random variable a satisfies the monotone hazard rate condition. \square