<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Sequential pricing for social networks with multi-state diffusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Niu, G; Li, VOK; Long, Y</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>The 2013 IEEE Global Communications Conference (GLOBECOM 2013), Atlanta, GA., USA, 9-13 December 2013. In Globecom. IEEE Conference and Exhibition, 2013, p. 3176-3181</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2013</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/191600">http://hdl.handle.net/10722/191600</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>Globecom. IEEE Conference and Exhibition. Copyright © IEEE.; ©2013 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
Sequential Pricing For Social Networks With Multi-State Diffusion

Guolin Niu, Victor O.K. Li, Yi Long
Department of Electrical and Electronic Engineering
The University of Hong Kong
Pokfulam Road, Hong Kong, China
Email: {gniu, vli, yilong}@eee.hku.hk

Abstract—The rapid development of online social networks (OSN) makes viral marketing through the word-of-mouth effect possible. Designing effective marketing strategy is critical in monetizing the social networks. However, most existing studies focus on conducting effective influence maximization analysis to propagate information widely instead of explicitly incorporating the pricing factor to design intelligent marketing strategies. In this paper, we study the sequential pricing models in which a monopoly seller iteratively posts a sequence of public unique prices for a product in stages, and any interested buyer can buy the product at the posted price at a particular stage if his valuation is above the posted price. Specifically, our model is built on the multi-state diffusion scheme where an active user may be in the “AWARE” and “ADOPT” states. Users in “ADOPT” state could influence their neighbors’ valuation for the given product. To realize the goal of revenue maximization, we develop a Dynamic Programming Based Heuristic (DPBH) to obtain the optimal pricing sequence. Application of the DPBH in the revenue maximization problem shows that it performs well in both the expected revenue achieved and in running time. This leads to fundamental ramifications to many related OSN marketing applications.

Keywords—Pricing, Revenue Maximization, Marketing, Monetizing Online Social Network.

I. INTRODUCTION

With the rapid development of online social networks (OSN) such as Facebook, Twitter and Google+, large-scale, instantaneous information dissemination becomes possible. This offers the potential for a surge of innovations and opportunities in social advertising and viral marketing [1]. However, most of the existing studies focus on conducting effective influence maximization analysis to propagate information widely instead of explicitly incorporating the pricing factor to design intelligent marketing strategies. In other words, extracting the economic value of social networks is still lacking [2]. Suppose a seller wants to sell a digital product via a social network, with intelligent selling strategies, i.e., pricing schemes, the revenue could be maximized. This is one way to monetize the social network, and is the research problem we will address in this paper.

There is a series of studies on viral marketing in the context of social networks. Domingos [3] and Richardson [4] are the first to formulate this as an algorithmic problem. They studied how to effectively promote a new technology or a new idea in an OSN by targeting a group of influential users. The seminal paper by Kempe et al. [5] then reformulated the problem further as “influence maximization” (INFMAX) utilizing discrete optimization theory. Given a social graph, and a positive number $k$, INFMAX is defined as finding a seed-set of $k$ nodes such that by targeting them initially, the viral spread is maximized.

While the INFMAX problem has been studied intensively, especially in the spreading of information and free products, it does not discuss the dependence of adoption on price. In economics, the pricing factor has been considered fully. However, the network effects are always addressed in a global or macro view. A recent stream of literature in computer science starts to study the “Revenue Maximization” (REVMAX) problem over social networks. The pioneering work by Hartline et al. [6] and Dsmitris et al. [7] studied optimal marketing for digital goods in OSNs and proposed the “Influence and Exploit” (IE) strategy. In IE, seed nodes are offered products free of charge, and then the seller could approach other users individually to offer specific prices. However, price discrimination, although useful for revenue maximization in some settings, may result in negative reactions from buyers. For example Oliver and Shor [8] [9] suggest that price discrimination could reduce the likelihood of purchasing. Moreover, it is impractical for sellers to adjust prices too frequently. Hence, we will reformulate the REVMAX problem such that price discrimination is prohibited.

While many studies have been conducted on REVMAX, most existing work does not distinguish the different types of active node states by simply assuming that active nodes will definitely adopt the given product. Researchers from different disciplines have demonstrated the differences between awareness and adoption. In sociology, Rogers (1962) [10] proposed five stages of product adoption including knowledge, persuasion, decision, implementation, and confirmation. Besides, in management science, Kalish [11] also characterized the adoption of a new product as consisting of two stages, namely, awareness and adoption. In computer science, simulations using real world datasets [12] also show that diffusion models which utilize “influenced state” and “adoption state” to depict “active” could reflect reality well. In particular, [13] proposed a profit maximization algorithm under multi-state diffusion settings. However, they still follow the “IE” framework, i.e., price discrimination still exists. Moreover, in their work sequential pricing scheme is not addressed.

To address the aforementioned limitations, we propose to study the problem of revenue maximization over online social networks, by incorporating a multi-state diffusion scheme. In particular, we investigate marketing strategies to maximize...
revenue of durable digital goods, where the cost of producing a copy of the goods is negligible and a customer is only interested in buying a single copy. By designing an optimal sequential pricing policy based on dynamic programming, the maximum revenue could be achieved within reasonable time. Our contributions could be summarized as follows.

- We propose a multi-state diffusion scheme by incorporating both the awareness propagation and the adoption diffusion. Here the awareness information is determined by social network effects and external influence (like mass media effects), while the actual adoption depends on the product price information and buyer’s valuation of the product.

- We formally formulate the sequential pricing problem over OSN with a multi-state diffusion scheme. Given $K$ possible stages, we determine a price vector $p$ with $K$ sequential prices to maximize the revenue. Recently, Akhlaghpour et al. [14] also studied the iterative pricing problem, but they did not consider the comprehensive diffusion process.

- We design an optimal pricing algorithm for any given $K$ via dynamic programming.

- We conduct experiments using two real world datasets to evaluate our proposed algorithm. Experimental results show that the algorithm performs well in the expected revenue achieved and in running time.

The rest of the paper is organized as follows. Section II discusses the proposed framework including the multi-state diffusion scheme and the REVMAX problem formulation. Section III presents our pricing strategies. Experiments are discussed in Section IV. Section V provides conclusions and future work.

II. PROPOSED FRAMEWORK

In this section we will firstly introduce the proposed multi-state diffusion model (MSDM) and then formally formulate the REVMAX problem. The notations are shown in Table I.

A. Diffusion Model

In this part, we introduce our multi-state diffusion scheme. The social graph can be represented in the form of a directed graph $G=(V,E)$, where $V = \{u_1, u_2, ..., u_N\}$ is the set of nodes (users), and $u_i$ is the unique ID of User $i$, $E = \{(u, v) | u, v \in V\}$ is the set of edges. Each user belongs to one of three states: INACTIVE (User is not aware of the product), AWARE (User is aware or develops some interest about the product), and ADOPT (User decides to adopt the product). The state transition diagram is shown in Fig. 1. Here, $g_i(X, Y, Z)$ is the influence function for $u_i$, and it dictates $u_i$’s awareness decision. $W_i(X_i(t))$ is $u_i$’s valuation function, and it depends on the set of $u_i$’s friends who has already bought the product. Here, we assume that the users in our system are fully aware of their friends’ state information. More details will be specified in the following subsections.

1) Awareness Diffusion: Inactive nodes become AWARE due to word-of-mouth effect or advertising. Thus the likelihood of becoming AWARE is proportional to the number of “transmitters” in the social network (i.e. $X(t)$, $Y(t)$, $Z(t)$) and the exogenous advertising effects (i.e. $A(t)$). See Table I for the notations. Fig. 2 shows how different types of users influence the target user.

For a given user $u_i$, the awareness diffusion probability equation is shown as follows.

$$g_i(X, Y, Z) = A(t) + \alpha \frac{|X_i(t)|}{|N_i|} + \beta \frac{|Y_i(t)|}{|N_i|} + \gamma \frac{|Z_i(t)|}{|N_i|} \quad (1)$$

$$|X_i(t)| = |X(t) \cap N_i| \quad (2)$$

$$|Y_i(t)| = |Y(t) \cap N_i| \quad (3)$$

$$|Z_i(t)| = |Z(t) \cap N_i| \quad (4)$$

This awareness equation says that, an inactive user $u_i$ will become AWARE with probability $g_i(X, Y, Z)$, and remains INACTIVE with probability $1 - g_i(X, Y, Z)$. $A(t)$ is the external influence, and other terms represent the influence exerted by the neighbors of $u_i$. Here, we make a close-world assumption that adopters transmit information more effectively than AWARE individuals, i.e., $\alpha > \beta > 0$, and INACTIVE nodes do not have influential power, i.e., $\gamma = 0$. 

![Fig. 1. The State Transition Diagram](image)

![Fig. 2. The Awareness Diffusion](image)

### TABLE I. Notations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_i$</td>
<td>Set of $u_i$’s friends.</td>
</tr>
<tr>
<td>$X_i(t)$</td>
<td>Initial set of users in ADOPT state at the beginning of pricing stage $t$</td>
</tr>
<tr>
<td>$X_e(t)$</td>
<td>Initial set of users in ADOPT state at the end of pricing stage $t$</td>
</tr>
<tr>
<td>$Y_i(t)$</td>
<td>Final set of users in AWARE state at the beginning of pricing stage $t$</td>
</tr>
<tr>
<td>$Y_e(t)$</td>
<td>Final set of users in AWARE state at the end of pricing stage $t$</td>
</tr>
<tr>
<td>$Z_i(t)$</td>
<td>Initial set of users in INACTIVE state at the beginning of pricing stage $t$</td>
</tr>
<tr>
<td>$Z_e(t)$</td>
<td>Final set of users in INACTIVE state at the end of pricing stage $t$</td>
</tr>
<tr>
<td>$A(t)$</td>
<td>Likelihood that any user is exposed to the external advertising at pricing stage $t$</td>
</tr>
<tr>
<td>$p_t$</td>
<td>Posted price of the given product at pricing stage $t$</td>
</tr>
<tr>
<td>$K$</td>
<td>Total number of pricing stages</td>
</tr>
<tr>
<td>$W_i(X_i(t))$</td>
<td>User $u_i$’s valuation of the product, and it depends on the users who already own the product, $</td>
</tr>
</tbody>
</table>
2) Adoption Diffusion: This section introduces the adoption diffusion scheme. Due to personal differences or human diversities, different people will value the same product differently. For example, the income, personal interests and actual needs will affect the adoption process. Existing literature [15] uses the term “reservation prices” to represent user valuation. A reservation price is the maximum amount a customer would be willing to pay given current available information. If the actual price is below the individual’s reservation price, he will buy; otherwise, he will not. As more users adopt the product, the valuation of the product increases. At stage $t$, $W_I(X_i(t))$ represents $u_i$’s valuation of the product.

$$W_I(X_i(t)) = I_i + \Psi_I(X(t) \cap N_i) \quad (5)$$

Note that $I_i$ represents $u_i$’s intrinsic value in buying certain products and $\Psi_I(X_i(t))$ is purely determined by the set $X_i(t)$. So we have $\Psi_I(0) = 0, \Psi_I(0) = I_i$. It means that even if no friends of $u_i$ buy the given product, $u_i$’s valuation is still not zero, since a customer may have certain intrinsic value for a given product.

If the current announced price is $p_i$, $u_i$ will adopt the given product under the condition that $W_I(X_i(t)) \geq p_i$. Here, we make the assumption that the seller does not know the buyer’s exact valuation, but it does know the distribution $V_i$ from which the valuation is drawn. Moreover, we assume that users have positive externalities on each other, i.e., for two subsets, $S, T \subseteq V$, if $S \subseteq T \subseteq V$, then $W_I(S) < W_I(T)$.

3) Diffusion Dynamics with Sequential Pricing: Instead of just publishing one unique price at all times, we allow the seller to post $K$ public prices sequentially to promote the given product. Each pricing stage ends when no more buyers are willing to buy the product at the current price. Then the seller has to lower the price to the current maximum valuation among the AWARE nodes, and start the next pricing stage. The K-stage sequential pricing process ends if either of the following two conditions occurs: 1) The price has been posted for K stages; 2) The number of nodes in AWARE state is 0. Obviously, following the proposed diffusion framework, the sequence of prices should be decreasing. The diffusion dynamics within a given pricing stage $t$ is as follows:

- Find $u_i$, the node with maximum valuation within the AWARE set. Set the public posted price for the current stage $p_t = W_i$ and set $u_i$’s state to ADOPT.
- Update the states for $u_i$’s existing neighbours using Algorithm 1 (Node Decision Algorithm).
- If newly ADOPT nodes or AWARE nodes are found, continue the diffusion process.
- The current pricing stage $t$ ends if no more state transition is possible.

Algorithm 1 Node Decision Algorithm

```
INPUT $u_i$’s current state $s_t$, current price $p_t$
if $s_t$ is INACTIVE
  $u_i$ becomes AWARE with probability $g_i$
else if $W_i > p_t$ then
  set $s_t$ to ADOPT
else
  set $s_t$ to AWARE
```

When the buyer accepts the offered price, i.e., the price of the product is smaller than the buyer’s valuation, the seller earns the price of the product as the revenue. Suppose the price at stage $t$ is $p_t$, and the corresponding number of buyers is $B_t$, then the system revenue at pricing stage $t$, i.e., $R_t$, will be:

$$R(t) = p_t \times B_t \quad (6)$$

Consider that in the whole marketing period, the system could post at most $K$ different prices, i.e., there are $K$ pricing stages. Then our goal is to find a price vector $p = \{p_1, p_2, ... , p_K\}$ of $K$ prices to maximize the cumulative revenue $\sum_{0 \leq t \leq K} R_t(p_t)$, then the optimal pricing sequence is:

$$p = \arg \max_{0 \leq t \leq K} \sum_{0 \leq t \leq K} R_t = \alpha \arg \max_{0 \leq t \leq K} \sum_{0 \leq t \leq K} p_t B_t \quad (7)$$

This is defined as the REVMAX(K) problem.

III. METHODOLOGY

In this section, we first study a restricted case of the REVMAX(K) problem, and show that we can identify the optimal pricing strategy based on a simple dynamic programming approach. For the general case, due to the complexity of the problem it is hard to find an optimal solution. Hence we propose a Dynamic Programming Based Heuristic (DPBH) to achieve a good approximation for REVMAX(K) within polynomial time.

A. A Special Case of REVMAX(K)

To better understand the properties of REVMAX(K), we first study a special case of the problem. We assume that a user’s valuation of the given product is deterministic, i.e., for each $u_i$, we have $W_I(X_i) = W_i$. For simplicity, the awareness diffusion probability $g_i(X, Y, Z)$ is generated before the diffusion starts. Under this setting, we generate a pricing vector $p = \{p_1, p_2, ... , p_T\}$, $p_1 > p_2 > ... > p_T$ following the greedy algorithm (Algorithm 2). Here, $T$ is the total number of pricing stages required to get all potential users to adopt.

The system revenue here is calculated as

$$\text{revenue} = p_1 X_{p_1} + p_2 X_{p_1 p_2} + ... + p_T X_{p_1 p_2 ... p_T} \quad (8)$$
Algorithm 2 Pricing Sequence Generation

Set $t=1$, then $p_t = \text{max}(W_i)$, for each $u_i$ in $V$
Add $p_t$ to vector $\mathbf{p}$
update node states in node set $V$ according to Algorithm 1
while AWARE node set is not empty do
  set $p_t = \text{max}(W_i)$, for each $u_i$ in AWARE set
  add $p_t$ to vector $\mathbf{p}$
  update node states in node set $V$
end while
return vector $\mathbf{p}$

Algorithm 3 Candidate Price Set Construction

set $t_{\text{min}} = \left\lfloor \log_{1+\varepsilon} \text{opt}_K \right\rfloor$, $t_{\text{max}} = \left\lceil \log_{1+\varepsilon} \text{opt}_K \right\rceil$
for $i = t_{\text{min}}$ to $t_{\text{max}}$ do
  $p_i = (1+\varepsilon)^{t_{\text{max}}-t_{\text{min}}-i}$
end for
return $p$

TABLE II. STATISTICS OF THE TWO REAL-WORLD NETWORKS

<table>
<thead>
<tr>
<th>Dataset</th>
<th>arXiv HEP-PH</th>
<th>DBLP community</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodes</td>
<td>11,204</td>
<td>1,772</td>
</tr>
<tr>
<td>Edges</td>
<td>235,238</td>
<td>14,984</td>
</tr>
<tr>
<td>Graph density</td>
<td>0.0019</td>
<td>0.0048</td>
</tr>
</tbody>
</table>

Now we consider the general REVMAX($K$) problem.

While it is difficult to get an optimal solution, using heuristics, we could get a polynomial time approximation regardless of the specific form of the valuation distribution. Our proposed Dynamic Programming Based Heuristics (DPBH) consists of the following four steps:

* Specify the price bounds $p_{\text{min}}$ and $p_{\text{max}}$, ensuring that the optimal pricing sequence $\mathbf{p} = \{p_1, p_2, \ldots, p_K\}$ satisfies $p_i \in [p_{\text{min}}, p_{\text{max}}]$, $i = 1, 2, \ldots, K$.
* Construct candidate price set using Algorithm 3.
* Simulate the diffusion process for sufficient runs to get the expected buyer number for all prices within the candidate price set.
* Use dynamic programming to choose the optimal pricing sequence based on the optimal substructure:

$$ R[K, p] = \max_{0 < p' < p} \{ R[K-1, p'] + p'(X_p' - X_p) \} $$

Let the size of the candidate price vector generated by Algorithm 3 be $M$. $M = \left\lceil \log_{1+\varepsilon} \text{opt}_K \right\rceil$, and the computational complexity of our proposed DPBH is $K M = K \left\lceil \log_{1+\varepsilon} \text{opt}_K \right\rceil$, where $K$ is the number of pricing stages.

Consider the performance of the proposed heuristic.

Under certain conditions, in which the random variable $X_p$ is i.i.d, then the maximum revenue generated by the DPBH, $E[\text{rev}(K)]$, can provide a performance guarantee within a factor of $1 - \frac{\varepsilon}{(1+\varepsilon)^2}$ of the optimal revenue, $E[\text{rev}(K)]$.

The sketch of the proof is as follows. Since $X_p$ is i.i.d and $0 < X_p < \left| V \right|$, using Chernoff-Hoeffding theorem, we can show that with high probability $(1 - \varepsilon)E[X_p] \leq X_p \leq (1 + \varepsilon)E[X_p]$. Suppose $u_i$ buys the product at price $(1 + \varepsilon)^r < p_{\text{opt}} < (1 + \varepsilon)^{r+1}$ in the optimal solution, then $u_i$ will buy the product when the price is $(1 + \varepsilon)^r$ in our solution, and we have $(1 + \varepsilon)^r > p_{\text{opt}}$. Combining the above approximations of the adopt number and the candidate price set construction, we show that the maximum revenue achieved using our heuristic is within $1 - \frac{\varepsilon}{(1+\varepsilon)^2}$ of the optimum.
network, is from the e-print arXiv and covers scientific collaborations submitted to HEP-PH category. The edges represent the collaboration relationships between authors. Here we use the largest connected component of the arXiv HEP-PH dataset. The second graph, the DBLP computer science bibliography provides a comprehensive list of research papers in computer science. Here we just focus on one comparatively denser community of the DBLP graph. The statistics for both graphs are shown in Table II. Here, the edges are directed, so the density metric is calculated as $|E|/|V|(|V| - 1))$.

**Algorithms.** Here we compare our proposed DPBH with the baseline random algorithm which chooses the pricing sequence randomly. For each given input, 1000 simulation runs are used to obtain an accurate estimate of the adoption spread. For all buyers, we adopt Gaussian distribution and Uniform distribution to model user’s valuation behaviour. The Gaussian distribution has been justified using real datasets and utilized in existing papers [13] [17]. For unfamiliar products, the Uniform distribution is used to account for lack of knowledge [18] [19].

**B. Experimental Results**

In this section we show the experimental results of four metrics. To save space, we use Fig 3 to illustrate the first three metrics.

**The Optimal Pricing Sequence.** Fig. 3(a) and Fig. 3(d) show the optimal sequence for $K = 50$ and $K = 100$, respectively. As shown in both figures, the decreasing trend of the price sequence will be affected by the valuation function. For the Uniform distribution, it drops linearly. For the Gaussian distribution, the trend is comparatively less steep.

**Revenue vs Number of Pricing Stages.** Fig. 3(b) and Fig. 3(e) study the increase of maximized revenue as the number of pricing stages increases, under the Gaussian and Uniform distributions, respectively. In both cases, the maximum achievable revenue increases with the increase of pricing stages. However, there is an obvious diminishing returns phenomenon, i.e., within a limited number of pricing stages (around 20 for DPBH, and 30 for the random algorithm), the seller could extract almost the same revenue that is achieved by more pricing stages.

**Adopt Number vs Number of Pricing Stages.** Fig. 3(c) and Fig. 3(f) study the increase of adopt number as the number of pricing stages increases under the Gaussian and Uniform distributions, respectively. The results show that maximized influence (adopt number) does not necessarily mean maximized revenue. Under the Gaussian distribution, the random algorithm performs equally well as DPBH for the adopt number metric. Under the Uniform distribution, while DPBH performs much better than the random algorithm in terms of the revenue, the random algorithm outperforms DPBH in terms of the adopt number.

**Graph Density vs Pricing Strategy.** Finally, we study how the network topology may influence the pricing strategy. Here we use two real world networks with different graph densities. Fig. 4 and Fig. 5 show how the pricing strategy changes under different network structures, for the Gaussian and the Uniform distributions, respectively. Here, the DBLP network is almost three times denser than the arXiv network. Although the magnitudes of the maximum achievable revenue are different due to the differences in the network scale, the diminishing returns phenomenon still holds for both networks, for both the Gaussian and the Uniform distributions.
We study the problem of revenue maximization over online social networks, by incorporating the multi-state diffusion scheme. In particular, we investigate marketing strategies to maximize the revenue of durable digital goods. By designing a sequential optimal pricing algorithm using dynamic programming, the maximum achievable revenue could be obtained within polynomial time. Moreover, our algorithm could provide performance guarantees under specific cases. Experimental results using real-world datasets demonstrate the effectiveness of our proposed heuristic.

In our current studies, the cost of the product and repeat purchase effect are not considered. We will consider these issues in our future work.

V. CONCLUSION

We study the problem of revenue maximization over online social networks, by incorporating the multi-state diffusion scheme. In particular, we investigate marketing strategies to maximize the revenue of durable digital goods. By designing a sequential optimal pricing algorithm using dynamic programming, the maximum achievable revenue could be obtained within polynomial time. Moreover, our algorithm could provide performance guarantees under specific cases. Experimental results using real-world datasets demonstrate the effectiveness of our proposed heuristic.

In our current studies, the cost of the product and repeat purchase effect are not considered. We will consider these issues in our future work.

REFERENCES


