Budget Allocation Problem with Multiple Advertisers: A Game Theoretic View

Takanori Maehara[1,4] MAEHARA.TAKANORI@SHIZUOKA.AC.JP
Akihiro Yabe[2,4] A-YABE@CQ.JP.NEC.COM
Ken-ichi Kawarabayashi[3,4] K_KENITI@NII.AC.JP
[1]Shizuoka University, [2]NEC Corporation, [3]National Institute of Informatics
[4]JST, ERATO, Kawarabayashi Large Graph Project

Abstract

In marketing planning, advertisers seek to maximize the number of customers by allocating given budgets to each media channel effectively. The budget allocation problem with a bipartite influence model captures this scenario; however, the model is problematic because it assumes there is only one advertiser in the market. In reality, there are many advertisers which are in conflict of advertisement; thus we must extend the model for such a case.

By extending the budget allocation problem with a bipartite influence model, we propose a game-theoretic model problem that considers many advertisers. By simulating our model, we can analyze the behavior of a media channel market, e.g., we can estimate which media channels are allocated by an advertiser, and which customers are influenced by an advertiser.

Our model has many attractive features. First, our model is a potential game; therefore, it has a pure Nash equilibrium. Second, any Nash equilibrium of our game has 2-optimal social utility, i.e., the price of anarchy is 2. Finally, the proposed model can be simulated very efficiently; thus it can be used to analyze large markets.

1. Introduction

1.1. Background

Marketing is used by *advertisers* to create loyal *customers*. A major decision problem in marketing planning is the allocation of budgets to *media channels* (e.g., TV, newspaper, billboards, and websites) efficiently to maximize the number of influenced customers under some constraints. This problem can be formulated as follows: Suppose we have estimates for the extent to which marketing channels can influence customer decisions and convert potential customers into loyal buyers. We would like to market a new product that may possibly be adopted by a large fraction of the customers. How should we choose the appropriate marketing channels?

Alon et al. (Alon et al., 2012) were the first to model this channel market. Consider a bipartite graph in which one side is the set of media channels and the other is the customers. Each edge between a channel and a customer indicates that the customer is influenced via the channel with some probability that depends on the budget allocated to that channel. The maximization of the expected number of influenced customers under some budget constraint can be formulated as a combinatorial optimization problem. This optimization problem can be developed in the general *submodularity* framework. More precisely, it can be formulated as *knapsack-constrained monotone submodular function maximization problem on the integer lattice*. Thus a polynomial time $(1 - 1/e)$-approximation algorithm can be obtained by a greedy type algorithm (Alon et al., 2012; Soma et al., 2014). We describe this submodularity framework in more details in Section 3. This model and its extensions are studied in the area of computational advertising (Ieong et al., 2014; Geyik et al., 2014; Hatano et al., 2015).

1.2. Motivation

In reality, there are *many advertisers* with comparable products in a channel market, and they *compete* to convert potential customers into as many loyal buyers as possible.

Market analysts (often controlled by *publishers* or *advertisers*) analyze such a complicated market. Major tasks of analysts are to simulate/predict the market to acquire an insight, and to propose a strategy for a market planning (Broder & Josifovski, 2011). In particular, they wish to estimate the media channels that are allocated by advertisers and the number of customers that are influenced by

Proceedings of the 32^{nd} *International Conference on Machine Learning*, Lille, France, 2015. JMLR: W&CP volume 37.

each advertiser. Since the model proposed in (Alon et al., 2012) only deals with a single advertiser, analysts need a market model that deals with many advertisers. This is the main motivation of our paper.

As mentioned above, many advertisers are competing, and they want to convert as many potential customers into loyal buyers as possible via media channels. In this case, each advertiser would adopt his own strategy independent of his competitors' marketing strategies. Thus we need to develop a network equilibrium framework that considers the reality of competition among multiple advertisers through media channels.

For this purpose, it makes sense to adopt *game-theoretic analysis*; we regard a media channel market as a *game* in which advertisers act as the *players* of the game. The goal of each player is to maximize his *utility* (the number of influenced customers). Here, we assume that each player knows the influence probability of all other players, i.e., we here consider a *complete information game*. This assumption is reasonable when there are sufficiently many transaction logs of the market to construct models of players. Moreover, we here assume that each player is *rational and selfish*, i.e., he chooses a strategy that (approximately) maximizes his utility. This assumption is, in particular, reasonable for online advertising and internet auctions (Varian, 2007; Zhao & Nagurney, 2008) since many advertisers use computer programs in these markets (Yuan et al., 2013). In this setting, game theory provides a rich theory for analyzing a game.

1.3. Contribution

We propose a game-theoretic model of a media channel market with n advertisers and discuss its properties. We offer the following theoretical contributions.

- We propose a game-theoretic model of a channel media market with n advertisers, which is an extension of the budget allocation problem presented in (Alon et al., 2012). We call this model the *budget allocation game with bipartite influence model* (Section 4).
- We show that the game has a *pure Nash equilibrium*. This is a desirable property for a market model. To prove this property, we show that the game is a *potential game* (Section 5).
- The social utility of any pure Nash equilibrium is 2-optimal. Furthermore, there exists a Nash equilibrium of $\min\{H_n, 2\}$-optimal, where $H_n = 1 + 1/2 + \cdots + 1/n$ is the n-th harmonic number. We give an asymptotically tight example for these results (Section 6).

To show the practical usefulness of our model, we also discuss the computational perspective.

- Finding a pure Nash equilibrium is NP-hard. Thus we consider an *approximate Nash equilibrium* and prove that it can be found in polynomial time. We also give a bound of the approximate price of anarchy (Section 7).
- We conduct numerical experiments with synthetic and real datasets. We observe that the approximate best-response dynamics converges in only a few rounds, and the social utility of the obtained solution is very close to the social optimal. These are much better than theory. Thus, our model has a capability of analyzing large markets, which will appear in online advertising and internet auctions (Section 8).

All omitted proofs will be given in Appendix A in the supplementary material.

2. Notation

In this section, we first introduce game theoretical notions and submodularity.

2.1. Game theory

We first define a (non-cooperative) game. Let $n \in \mathbb{Z}_{>0}$ be the number of players and let $[n] = \{1, \ldots, n\}$ be the set of players. An *n-player game in strategic form (Nisan et al., 2007) (a "game" for short)* is a tuple $([n], \{\mathcal{D}_i\}_{i\in[n]}, \{f_i\}_{i\in[n]})$, where $\mathcal{D}_i$ is a set of *strategies* for player i, and $f_i : \mathcal{D} \to \mathbb{R}$ is a utility function of player i, where $\mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ is the set of all possible *configurations*. The goal of player i is to find a strategy $x_i \in \mathcal{D}_i$ to maximizes his utility function f_i. For notational convenience, when there is no fear of confusion, we denote by $g(x_i, x_{-i})$ the value of function $g : \mathcal{D} \to \mathbb{R}$ when player i selects a strategy x_i and other players select strategies $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$.

A configuration $x \in \mathcal{D}$ is a (pure) *Nash equilibrium* (Nash, 1950) if it satisfies

$$f_i(x_i, x_{-i}) \geq f_i(x'_i, x_{-i}) \tag{1}$$

for all $x'_i \in \mathcal{D}_i$. Nash equilibrium is the most commonly-used notion of equilibrium in game theory, and in general, a game does not necessarily have a pure Nash equilibrium. See (Nisan et al., 2007) for more details of algorithmic game theory.

2.2. Submodular function on integer lattice

Let S be a finite set. A function $f : \mathbb{Z}^S \to \mathbb{R}$ is *submodular on the integer lattice ("submodular")* if it satisfies

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{2}$$

for all $x, y \in \mathbb{Z}^S$ where $\vee$ denotes the element-wise maximum and $\wedge$ denotes the element-wise minimum:

$(x \vee y)(s) = \max\{x(s), y(s)\}$ and $(x \wedge y)(s) = \min\{x(s), y(s)\}$. A function $f : \mathbb{Z}^S \to \mathbb{R}$ is *monotone nondecreasing* if it satisfies

$$f(x) \leq f(y) \tag{3}$$

for all $x \leq y$, i.e., $x(s) \leq y(s)$ ($s \in S$). A function $f : \mathbb{Z}^S \to \mathbb{R}$ satisfies *component-wise concavity* (Shioura, 2009; Milgrom & Strulovici, 2009) if it satisfies

$$f(x + e_s) - f(x) \geq f(x + 2e_s) - f(x + e_s) \tag{4}$$

for all $s \in S$, where e_s is the s-th unit vector.[1] See (Topkis, 1998) for more details of submodular function defined on the lattice.

3. Budget allocation problem

Here, we review the budget allocation problem with a bipartite influence model, which is proposed by Alon et al. (Alon et al., 2012) and extended by Soma et al. (Soma et al., 2014). We discuss a game-theoretic extension of this model in the next section.

Let $G = (S \cup T, E)$ be a bipartite graph, where the left vertices S denote media channels, the right vertices T denote customers, and the edges $E \subseteq S \times T$ denote the relations between channels and customers. Each edge $(s,t) \in E$ has an *activation probability* $p(s,t) \in [0,1]$ such that customer $t \in T$ is *activated* via channel $s \in S$ with probability $p(s,t)$.

Suppose that an advertiser has a budget of $B \in \mathbb{Z}_{\geq 0}$ units. If $x(s) \in \mathbb{Z}_{\geq 0}$ units are allocated for channels $s \in S$, then customer $t \in T$ is activated by the advertiser with probability

$$P(x,t) = 1 - \prod_{s \in \Gamma(t)} (1 - p(s,t))^{x(s)}, \tag{5}$$

where $\Gamma(t) = \{s \in S : (s,t) \in E\}$ denotes the neighbor of $t \in T$. Thus, the *expected number of activated customers by the advertiser* $f(x)$ is given by

$$\begin{aligned} f(x) &= \sum_{t \in T} P(x,t) \\ &= \sum_{t \in T} \left(1 - \prod_{s \in \Gamma(t)} (1 - p(s,t))^{x(s)}\right). \end{aligned} \tag{6}$$

The purpose of advertiser is to maximize $f(x)$ under budget constraints, which is specified as follows. Let $c(s) \in \mathbb{Z}_{\geq 0}$ be a capacity of channel $s \in S$ and $w(s) \in \mathbb{Z}_{\geq 0}$ be an allocation cost of channel $s \in S$. These denote that the channel s can be allocated at most $c(s)$ units and the cost of allocating an unit is $w(s)$. The advertiser wants to allocate the budgets to each channel under the capacity and the cost constraints. Thus his/her problem is given as the following.

$$\begin{array}{ll} \text{maximize} & f(x) \\ \text{subject to} & \sum_{s \in S} w(s)x(s) \leq B, \\ & 0 \leq x(s) \leq c(s). \end{array} \tag{7}$$

The function f defined by (6) is a monotone submodular function on the integer lattice and satisfies component-wise concavity (Soma et al., 2014). For a general monotone submodular function f with component-wise concavity, the problem (7) is NP-hard. Moreover, it is NP-hard to obtain better than $(1 - 1/e)$-approximate solution of the problem (Alon et al., 2012). Thus we are interested in approximation algorithms.

Alon et al. proposed two polynomial time algorithms for the problem (7). The first algorithm is a $(1 - 1/e)$ approximation algorithm, which is an integer lattice version of Sviridenko (Sviridenko, 2004). This algorithm combines a *partial enumeration* and a *greedy procedure*; it first enumerates all feasible assignments of cardinality at most three, and then performs greedy procedure (Algorithm 1) from these assignments. Since the number of initial assignments is $O(|S|^3 B^3)$ and the greedy procedure terminates at most $|S|B$ iterations, the complexity of this algorithm is $O(\gamma |T||S|^4 B^4)$, where γ is the complexity of evaluating $P(x,t)$. Soma et al. (Soma et al., 2014) extended this algorithm for a general monotone submodular function on the integer lattice.

Even though the above algorithm runs in polynomial time, in practice, it is not scalable for large instances due to its expensive exponent. The second algorithm of Alon et al. (Alon et al., 2012) is a more efficient but less accurate algorithm. The algorithm first performs a greedy procedure (Algorithm 1) from empty assignment $x = 0$, and then compares the obtained solution with all feasible assignments of cardinality one. This procedure gives a $(1/2)(1 - 1/e)$ approximation solution in $O(\gamma |T||S|B)$ time. Note that this algorithm can be immediately extend to a general submodular function on the integer lattice with component-wise concavity.

4. Proposed model

In this section, we propose a game-theoretic extension of the budget allocation problem. Some possible extensions of the model are discussed in Section 9.

Suppose that there are n advertisers in a market. We here define a *game*, and regard these advertisers as the *players*

[1] This property is also called *diminishing marginal return property* (Soma et al., 2014). "Submodularity with component-wise concavity" is called *directional concavity* (Shaked & Shanthikumar, 1990), which was originally studied by Rüschendorf (Rüschendorf, 1983).

Algorithm 1 Greedy procedure for maximizing monotone submodular function with component-wise concavity under a knapsack constraint (Alon et al., 2012).

1: **loop**
2: $s^* \in \operatorname{argmax}_s\{(f(x+e_s)-f(x))/w(s) : x(s) \leq c(s)-1, \sum_{s' \in S} w(s')x(s') \leq B - w(s)\}$
3: **if** NO such s^* **then return** x as a solution
4: $x(s^*) \leftarrow x(s^*) + 1$
5: **end loop**

Algorithm 2 Activating customer t by the random ordering rule.

1: Draw a random permutation $\sigma \in \mathcal{S}_n$
2: **for** $k = 1, \ldots, n$ **do**
3: Player $\sigma(k)$ tries to activate customer t with probability $P_{\sigma(k)}(x_{\sigma(k)}, t)$
4: **if** The challenge is succeeded **then break**
5: **end for**

of the game. Similar to the budget allocation problem, we consider a bipartite graph $G = (S \cup T, E)$, where S denotes the set of channels and T denotes the set of customers. Each player $i \in [n]$ has activation probability $p_i : E \to [0, 1]$, budget $B_i \in \mathbb{Z}_{\geq 0}$, channel capacity $c_i : S \to \mathbb{Z}_{\geq 0}$, and allocation cost $w_i : S \to \mathbb{Z}_{\geq 0}$. The set of feasible strategies of player i is expressed as follows:

$$\mathcal{D}_i = \{x_i \in \mathbb{Z}^{|S|} : \sum_{s \in S} w_i(s)x_i(s) \leq B_i, \\ 0 \leq x_i(s) \leq c_i(s) \ (s \in S)\}. \quad (8)$$

The purpose of each player (advertiser) is to find a feasible allocation $x_i \in \mathcal{D}_i$ for maximizing the number of his customers.

Defining a game-theoretic extension of the budget allocation problem requires an activation rule, i.e., when many advertisers with comparable products simultaneously approach the same customer $t \in T$, we must specify which advertiser successfully activates that customer. We propose the following rule, which we refer to as the *random ordering rule*. First, we define the *contribution* of advertiser i as

$$P_i(x_i, t) = 1 - \prod_{s \in \Gamma(t)} (1 - p_i(s, t))^{x_i(s)} . \quad (9)$$

Note that (9) coincides with the activation probability when the other advertisers are ignored. For each customer $t \in T$, we draw a permutation $\sigma \in \mathcal{S}_n$ uniformly at random, where $\mathcal{S}_n$ denotes the set of all permutations of $[n]$. Then, according to permutation σ, each player sequentially attempts to activate customer t. The formal description is given in Algorithm 2.

The utility function $f_i(x_1, \ldots, x_n)$ of player $i \in [n]$ is the number of expected customers activated by player i when each player $j \in [n]$ allocates $x_j \in \mathcal{D}_j$. In the random ordering rule, we can obtain the the closed form of f_i as follows:

$$f_i(x) = \frac{1}{n!} \sum_{t \in T} \sum_{\sigma \in \mathcal{S}_n} P_i(x_i, t) \prod_{j \prec_\sigma i} (1 - P_j(x_j, t)), \quad (10)$$

where $i \prec_\sigma j$ indicates that i appears before j in σ.

The tuple $([n], \{f_i\}_{i \in [n]}, \{\mathcal{D}_i\}_{i \in [n]})$ defines a game. We call this the *budget allocation game with a bipartite influence model*. We propose this game for a model of a media channel market with many advertisers.

Intuitively, this model (random ordering rule) corresponds to the following customer's action: Imagine a customer who wants to buy a PC. Today he obtains many flyers from PC manufactures (advertisers) via media channels. He starts examining these manufactures sequentially, and when he identifies a desirable PC, he immediately purchases the PC and stops examining the other makers. If the examination order is random, the customer's actions correspond to the proposed model.

We here give a basic property of the utility functions.

Proposition 1. *Each utility function $f_i(x_i, x_{-i})$ is a monotone nondecreasing submodular function with component-wise concavity with respect to $x_i \in \mathcal{D}_i$.*

5. Existence of pure Nash equilibria

Existence of pure Nash equilibria is a desirable property of a market model, in particular, for market analysts. When a current configuration is not a Nash equilibrium, each player may change his strategy to improve his utility. Therefore, if market analysts want to propose a strategy for advertisers, a market model should be a Nash equilibrium.

Here, we show that a budget allocation game has a pure Nash equilibrium. To prove this, we introduce the notion of a potential game. A game is a *potential game* (Monderer & Shapley, 1996) if there exists a function $\Phi : \mathcal{D} \to \mathbb{R}$ such that for all $i \in [n]$,

$$f_i(x_i', x_{-i}) - f_i(x_i, x_{-i}) = \Phi(x_i', x_{-i}) - \Phi(x_i, x_{-i}). \quad (11)$$

Function Φ is called a *potential of the game*. The potential is unique up to a constant difference (Monderer & Shapley, 1996). The most important property of a potential game is that it must have a pure Nash equilibrium. In fact, by (11), the maximum potential feasible solution is a Nash equilibrium.

We show that our game is also a potential game; thus it has a pure Nash equilibrium.

Proposition 2. *A budget allocation game is a potential*

Algorithm 3 Best response dynamics.

1: Set arbitrary initial solutions $x_i \in \mathcal{D}_i$, $i \in [n]$
2: **repeat**
3: **for** $i \in [n]$ **do**
4: Compute $x_i \in \operatorname{argmax}\{f_i(y_i, x_{-i}) : y_i \in \mathcal{D}_i\}$
5: **if** $f_i(y_i, x_{-i}) > f_i(x_i, x_{-i})$ **then** Update $x_i \leftarrow y_i$
6: **end for**
7: **until** no update is performed

game with potential function

$$\Phi(x) = \sum_{t \in T} \sum_{\emptyset \neq I \subseteq [n]} \frac{(-1)^{|I|-1}}{|I|} \prod_{i \in I} P_i(x_i, t). \tag{12}$$

Corollary 3. *A budget allocation game has a pure Nash equilibrium.*

The proof of Proposition 2 will be given in Appendix A.

A Nash equilibrium of a potential game can be found by the best-response dynamics (Algorithm 3) in a finite number of iterations (Monderer & Shapley, 1996). This implies that, in our model, a market tends to be a Nash equilibrium.

6. Quality of Nash equilibrium

As discussed in the previous section, our budget allocation game has a pure Nash equilibrium. In this section, we analyze the *quality* of pure Nash equilibria. We define *social utility* as the sum of the utilities of all advertisers.

$$\begin{aligned} F(x) =& f_1(x) + \cdots + f_n(x) \\ =& \sum_{t \in T} \left(1 - \prod_{i=1}^{n} (1 - P_i(x_i, t)) \right). \end{aligned} \tag{13}$$

The social utility $F(x)$ is the expected number of customers activated by some advertiser. Market analysts sometimes desire a strategy for advertisers to obtain a large social utility. However, as mentioned in the previous section, advertisers would not accept a configuration that is not a Nash equilibrium. Therefore we are interested in the social utility at a Nash equilibrium. To measure this, we adopt the *price of anarchy* and the *price of stability*.

Price of Anarchy The ratio between the best social utility and the worst social utility under Nash equilibria is referred to as the *price of anarchy* (Koutsoupias & Papadimitriou, 1999), which is expressed as

$$\mathrm{PoA} = \frac{\max\{F(x) : x \in \mathcal{D}\}}{\min\{F(x) : x \text{ is a Nash}\}}.$$

Since a market tends to a Nash equilibrium, at least PoA-optimal customers are naturally activated. Thus, if PoA is close to one, we can say that the market is efficient.

Here we show that the price of anarchy of the budget allocation game is at most 2. To show this, we prove that our game is a monotone utility game. A game is a *monotone utility game on the integer lattice* if it satisfies the following three conditions.

1. $F(x)$ is a nondecreasing submodular function on the integer lattice, and satisfies component-wise concavity.
2. $F(x) \geq \sum_{i=1}^{n} f_i(x)$.[2]
3. $f_i(x) \geq F(x_i, x_{-i}) - F(0, x_{-i})$.

Note that utility games are originally defined on set functions (Vetta, 2002). Here we extend this concept to a game on the integer lattice. We first show that the price of anarchy of a monotone utility game is at most 2.

Proposition 4. *The price of anarchy of a monotone utility game on the integer lattice is at most* 2*.*

Proposition 4 is an integer lattice version of Theorem 5 presented by Vetta (Vetta, 2002). This proof (which will be given in Appendix A) requires component-wise concavity.

We can show that the budget allocation game is a monotone utility game on the integer lattice. Therefore, by Proposition 5 below (whose proof will be given in Appendix A), we can show that the price of anarchy of the game is at most 2.

Proposition 5. *The budget allocation game is a monotone utility game on the integer lattice.*

Corollary 6. *The price of anarchy of the budget allocation game is at most* 2*.*

Price of Stability There is another quality measure, called the *price of stability* (Anshelevich et al., 2008). The price of stability is the ratio between the best social utility and the best social utility under Nash equilibria:

$$\mathrm{PoS} = \frac{\max\{F(x) : x \in \mathcal{D}\}}{\max\{F(x) : x \text{ is a Nash}\}}.$$

When the price of stability is close to one, there is an acceptable configuration with possibly many activated customers. Thus market analysts can propose this configuration to advertisers to obtain a large social utility.

By definition, the price of stability is at most the price of anarchy. We can slightly improve this bound.

Proposition 7. *The price of stability of a budget allocation game is at most* $\max\{H_n, 2\}$*, where* H_n *is the harmonic number,* $H_n = 1 + 1/2 + \cdots + 1/n$*.*

[2] Since we defined social utility as the sum of the players' utility, condition 2) is satisfied. However, we define a basic utility game for a more general situation.

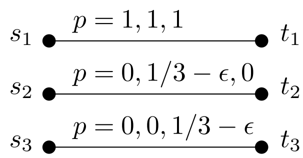

Figure 1. Asymptotically tight example for PoA and PoS.

Algorithm 4 Approximate best response dynamics.

1: Set arbitrary initial solutions $x_i \in \mathcal{D}_i$, $i \in [n]$
2: **repeat**
3: **for** $i \in [n]$ **do**
4: Find an η-approximate solution y_i of the problem $\max\{f_i(y_i, x_{-i}) : y_i \in \mathcal{D}_i\}$.
5: **if** $f_i(y_i, x_{-i}) > f_i(x_i, x_{-i}) + \epsilon$ **then** Update $x_i \leftarrow y_i$
6: **end for**
7: **until** no update is performed

Note that $H_2 = 1.5$, $H_3 \approx 1.833$, and $H_4 \approx 2.083$. Thus the above proposition (whose proof will be given in Appendix A) improves the bound for $n \leq 3$.

Example 8. *A bipartite graph $G = (S \cup T, E)$ is defined by n parallel lines, i.e., $S = [n]$, $T = [n]$, and $E = \{(i, i) : i \in [n]\}$. The edge probability is defined by*

$$p_i(j, j) = \begin{cases} 1, & j = 1, \\ 1/n - \epsilon, & j = i (\neq 1), \\ 0, & \textit{otherwise}. \end{cases} \tag{14}$$

Figure 1 shows the case for $n = 3$. Each $i \in S$ has unit cost and unbounded capacity, and each player has an unit budget.

In this instance, the social optimal solution is $x_i = e_i$ for all $i \in [n]$, and the corresponding social utility is $2 - 1/n$. On the other hand, a solution $x_i = e_1$ for all $i \in [n]$ is a unique Nash equilibrium. and the corresponding social utility is 1. Therefore PoA = PoS = $2 - 1/n$. By tending $n \to \infty$, we have PoA = PoS $\to 2$, i.e., it is a tight example for Propositions 6 and 7.

7. Approximate Nash equilibrium

In the previous sections, we considered only the theoretical perspective. Here, we consider the computational perspective.

Computing a Nash equilibrium is at least as hard as the budget allocation problem (7); thus it is NP-hard. Therefore we are interested in approximation algorithms. A configuration $x \in \mathcal{D}$ is an (η, ϵ)-*approximate Nash equilibrium* if it satisfies

$$f_i(x_i, x_{-i}) \geq \eta f_i(x'_i, x_{-i}) - \epsilon.$$

for all $x'_i \in \mathcal{D}_i$. For market analysis, an approximate Nash equilibrium has the following meaning: Since maximizing f_i is NP-hard, it makes sense to assume that

each advertiser adopts an (η, ϵ)-approximation algorithm to improve his strategy. (15)

In such a case, each advertiser may not change his strategy when the current strategy is an approximate Nash equilibrium, i.e., an approximate Nash equilibrium is an acceptable configuration for all advertisers.

We can find an approximate Nash equilibrium in polynomial number of iterations by *approximate best response dynamics* (Algorithm 4). This also implies that, in our model with (15), a market tends to be an approximate Nash equilibrium.

Proposition 9. *For an arbitrary $\epsilon > 0$, the best response dynamics finds an (η, ϵ)-approximate Nash equilibrium in $|T| H_n / \epsilon$ rounds (lines 2–9).*

If we use a polynomial time approximation algorithm for line 4, each round of the approximate best response dynamics requires polynomial number of utility functions evaluations. It should be mentioned that, in (10), the utility function f_i has $O(n!)$ terms. Thus the naive computation requires $O(n!)$ time. However, we can reduce this factor to $O(n(\log n)^2)$ by using a divide-and-conquer algorithm with the fast Fourier transform.

Proposition 10. *For each $i \in [n]$, $f_i(x)$ can be computed in $O(|E|n + |T|n(\log n)^2)$ time.*

If we adopt the $\eta = (1/2)(1 - 1/e)$ approximation algorithm (mentioned in Section 3) in line 4 of Algorithm 3, we can implement Algorithm 3 in $O(B|S||E|n + B|S||T|n(\log n)^2)$ time per round; see the proof of Proposition 10 in Appendix A.

Finally, we give a bound for the approximate price of anarchy. Again, the proof will be given in Appendix A.

Proposition 11. *Let $x^* \in \mathcal{D}$ be the optimal social utility allocation. For any (η, ϵ)-approximate Nash equilibrium $x \in \mathcal{D}$, we have*

$$F(x^*) \leq \left(1 + \frac{1}{\eta}\right) F(x) + \frac{\epsilon n}{\eta}.$$

This shows if each player uses $\eta = 1 - 1/e$ approximation algorithm for maximizing his utility, the approximate price of anarchy is about $1 + 1/(1 - 1/e) \approx 2.58$. If each player uses $\eta = (1/2)(1 - 1/e)$ approximation algorithm, the approximate price of anarchy is about $1 + 2/(1 - 1/e) \approx 4.16$.

8. Experiments

Here, we demonstrate the practical usefulness of the proposed model through numerical experiments.

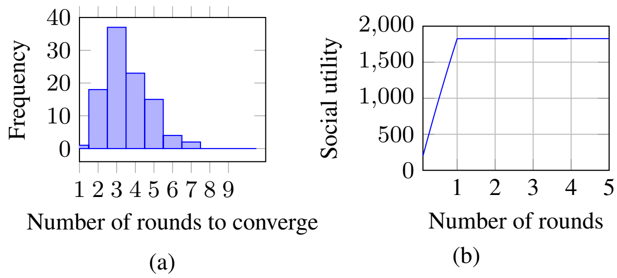

Figure 2. Number of rounds to converge.

Table 1. Computational time (per round) in real datasets.

dataset	$\lvert S\rvert$	$\lvert T\rvert$	n	time/round
open1	541	4271	20	24 s
open2	757	5062	20	40 s
movielens_100k	944	1683	10	3 s
movielens_1m	6040	3706	10	44 s
movielens_10m	69878	10677	10	1621 s
user-tag	17122	82035	5	518 s
user-picture	12177	495402	5	4728 s
tag-picture	82035	495402	5	18780 s

For efficient computation, we adopt the $(1/2)(1-1/e)$ approximation algorithm for line 4 in Algorithm 3, as mentioned right after Proposition 10. Furthermore, by using a naive algorithm, instead of the fast Fourier transform in evaluating f_i (see the proof of Proposition 10 in Appendix), the complexity per round becomes $O(B|S||T|n^2)$.

All experiments were conducted on an Intel Xeon E5-2690 2.90GHz CPU with 256GB memory running Ubuntu 12.04. Our algorithm was implemented in C++ and was compiled using g++v4.6 with the -O3 option.

8.1. Typical behavior

We first show the typical behavior of the algorithm. We generated 100 random bipartite graphs which has $|S| = 100$ and $|T| = 10000$ vertices, and each $t \in T$ has 20 random edges. We set $n = 10$ players, and choose each player's budgets and the influence probabilities such that $1/5$ of the customers will be activated. We performed our approximate best response dynamics with $\epsilon = 0.999$.

We observed that the dynamics converges in a few rounds. Figure 2 (a) shows the histogram of the number of rounds to converge. Most instances are converged in four rounds, and only a few instances require more than four founds (but at most seven rounds). To observe the precise behavior of the algorithm, we select one instance that requires five rounds and plot the social utility for each rounds. The result is shown in Figure 2 (b). This shows the process almost converges at the first round and it requires additional rounds for minor improvements.

We also performed similar experiments with other parameters and random graph models (e.g., random power-law graphs); however we obtain similar results for all experiments (see: Appendix B). Thus we conclude that this is a typical behavior of our model.

8.2. Scalability

We then discuss the scalability of our algorithm. We performed experiments for real datasets. We here adopt the following datasets.

1. *open-advertising-dataset*[3] (open1, open1). The dataset consists of click-query logs of a search engine.
2. *MovieLens dataset*[4] (movielens_100k, movielens_1m, and movielens_10m). MovieLens[5] is a web-based recommender system. The dataset consists of the rating of movies by users on MovieLens.
3. *vi.sualize.us dataset*[6] (user-tab, user-picture, and user-tag). vi.sualize.us[7] is a social bookmarking service for pictures. The dataset consists of bipartite graphs between two of users, tags, and pictures.

We construct bipartite graphs from these dataset. Budgets and influence probabilities of players are determined similarly as the experiment in Section 8.1. The results are summarized in Table 1. For each instance, the number of rounds for converge performs similar to the experiment in Section 8.1, i.e., the model converges in a few iterations. This shows our model can be efficiently simulated in large real datasets.

8.3. Approximate vs exact maximization

In our model, we assumed that each player approximately maximizes his utility. We here verify whether the exact maximization improves this utility.

Since the exact maximization requires exhaustive search, which is very expensive, we used a small instance for this experiment. We generated 100 random bipartite graphs with $|S| = 10$ and $|T| = 100$ vertices, and each $t \in T$ has 5 edges. We set $n = 2$ players. Other conditions are the same as the experiments in Section 8.1.

In this setting, the ratio between a player's utility when all players use the approximate maximization and when they use the exact maximization is contained in $[0.990, 1.008]$ (see Appendix B for more details). This means the approximate maximization is enough to evaluate the quality of so-

[3] https://code.google.com/p/open-advertising-dataset/
[4] http://grouplens.org/datasets/movielens/
[5] http://movielens.umn.edu/
[6] http://konect.uni-koblenz.de/networks/pics_ti
http://konect.uni-koblenz.de/networks/pics_ut
http://konect.uni-koblenz.de/networks/pics_ui
[7] http://vi.sualize.us/

lutions, which can be very efficiently computed.

8.4. Experimental value of price of anarchy

Finally, we evaluate the quality of the obtained social utility. Proposition 11 shows that the approximate price of anarchy is at most $1 + 1/\eta = 4.16$. We here compare this value with the social utility of the obtained solution. Here, we used the same instances in Section 8.3.

In this experiment, the ratio of the optimal social utility, obtained by exhaustive search and the obtained solution is at most 1.006, which is significantly better than the theory (see Appendix B). In fact, it is significantly better than the exact price of anarchy ($= 2$). This implies that, in our experiments, we can find an approximate Nash equilibrium which is very close to the social optimal solution.

9. Extension of the model

In this section, we discuss a possible generalization of our proposed model. In our theory, we have only used the property that $P_i(x_i, t)$ is a monotone nondecreasing submodular function with component-wise concavity. Thus, we can extend the proposed model for any probability function $P_i(X, t)$ that satisfies this property, instead of (9).

One important example is the *nonincreasing influence probability model* introduced by Soma et al. (Soma et al., 2014). In this model, each activation probability $p_i(e)$ is extended to nondecreasing probabilities $p_i^{(1)}(e) \geq p_i^{(2)}(e) \geq \cdots$, and $P_i(x_i, t)$ is extended to

$$P_i(x_i, t) := \prod_{s \in \Gamma(t)} \prod_{k=1}^{x(s)} (1 - p_i^{(k)}(s, t)). \tag{16}$$

This model captures real-world marketing phenomena, i.e., the effectiveness of activating a target customer may decrease in multiple trials.

10. Related work

This paper has discussed a game-theoretic extension of the budget allocation problem. To the best of our knowledge, this is the first game-theoretic extension of the problem. On the other hand, there are some game-theoretic models of the influence maximization problem (Kempe et al., 2003), which deals with information spreads over non-bipartite graphs. Here, we briefly review these models and discuss the relation with our model.

The first game-theoretic model of the influence maximization problem is proposed by Bharathi, Kempe, and Salek (Bharathi et al., 2007). In their model, each player initially selects seed vertices. Then the diffusion process is performed as same as the independent cascading model of the influence maximization problem. When a vertex u is simultaneously activated by many players, it belongs to one of the activated players. They proved that the price of anarchy of this model is at most 2.

Alon et al. (Alon et al., 2010) studied a model in which if many players simultaneously approach a vertex i, they "cancel out." They showed that, if the diameter of an underlying graph is at most 2 (which seems very restrictive), then the game has a pure Nash equilibrium. They also showed that the price of anarchy of their game is at most 2. Tzoumas and Amanatidis (Tzoumas et al., 2012) extended this model and discussed the condition of existence of a pure Nash equilibrium.

Goyal and Kearns (Goyal & Kearns, 2012) proposed a general influence model that is specified by a function, called a switching function. This is a generalization of the *linear threshold model* in the influence maximization problem (Kempe et al., 2003). He and Kempe (He & Kempe, 2013) analyzed this model and proved that the price of anarchy is at most 2.

By comparing these existing results, the largest advantage of our model is that it has a pure Nash equilibrium. As mentioned in Section 5, this is a desirable property for a market model.

11. Conclusion

In this paper we have proposed a model for a media channel market with many advertisers, which is a game-theoretic extension of the budget allocation problem presented by Alon et al. (Alon et al., 2012). The proposed model has the following attractive features: 1) it has a pure Nash equilibrium, 2) both the price of anarchy and the price of stability can be estimated, and 3) it can be efficiently simulated. We conducted numerical experiments to demonstrate that the model can be applied to a large scale market.

There are some possible future works. First, we here assumed that each player knows the influence probabilities of all other players. This makes our model a complete information game, which is easy to analyze. Extending this model to an incomplete information game is perhaps important future work. Second, we here assumed that an ordering of players is uniformly random, which is the key of Proposition 2. Introducing distributions over the orderings is a natural extension of the model; however, in this model, Proposition 2 does not hold, i.e., the existence of Nash equilibrium is non-trivial. Analyzing this extended model is interesting future work. Finally, in the experiments in Section 8.4, we have obtained an approximate Nash equilibrium which is very close to the social optimal solution. Exploiting the reason of this phenomenon seems practically important work.

References

Alon, Noga, Feldman, Michal, Procaccia, Ariel D, and Tennenholtz, Moshe. A note on competitive diffusion through social networks. *Information Processing Letters*, 110(6):221–225, 2010.

Alon, Noga, Gamzu, Iftah, and Tennenholtz, Moshe. Optimizing budget allocation among channels and influencers. In *Proceedings of the 21st International Conference on World Wide Web*, pp. 381–388, 2012.

Anshelevich, Elliot, Dasgupta, Anirban, Kleinberg, Jon, Tardos, Eva, Wexler, Tom, and Roughgarden, Tim. The price of stability for network design with fair cost allocation. *SIAM Journal on Computing*, 38(4):1602–1623, 2008.

Bharathi, Shishir, Kempe, David, and Salek, Mahyar. Competitive influence maximization in social networks. In *Internet and Network Economics*, pp. 306–311. 2007.

Broder, Andrei and Josifovski, Venja. MS&E 239: Introduction to Computational Advertising, 2011. URL `http://web.stanford.edu/class/msande239/`.

Geyik, Sahin Cem, Saxena, Abhishek, and Dasdan, Ali. Multi-touch attribution based budget allocation in online advertising. In *Proceedings of 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1–9, 2014.

Goyal, Sanjeev and Kearns, Michael. Competitive contagion in networks. In *Proceedings of the 44th annual ACM Symposium on Theory of Computing*, pp. 759–774, 2012.

Hatano, Daisuke, Fukunaga, Takuro, Maehara, Takanori, and Kawarabayashi, Ken-ichi. Lagrangian decomposition algorithm for allocating marketing channels. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, 2015.

He, Xinran and Kempe, David. Price of anarchy for the n-player competitive cascade game with submodular activation functions. In *Web and Internet Economics*, pp. 232–248. 2013.

Ieong, Samuel, Mahdian, Mohammad, and Vassilvitskii, Sergei. Advertising in a stream. In *Proceedings of the 23rd International Conference on World Wide Web*, pp. 29–38, 2014.

Kempe, David, Kleinberg, Jon, and Tardos, Éva. Maximizing the spread of influence through a social network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 137–146, 2003.

Koutsoupias, Elias and Papadimitriou, Christos. Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, pp. 404–413, 1999.

Li, Jian, Saha, Barna, and Deshpande, Amol. A unified approach to ranking in probabilistic databases. *Proceedings of the VLDB Endowment*, 2(1):502–513, 2009.

Milgrom, Paul and Strulovici, Bruno. Substitute goods, auctions, and equilibrium. *Journal of Economic Theory*, 144(1):212–247, 2009.

Monderer, Dov and Shapley, Lloyd S. Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258–265, 1996.

Nash, John. Equilibrium points in n-person games. *Proceedings of the National Academy of Science*, 36(1):48–49, 1950.

Nisan, Noam, Roughgarden, Tim, Tardos, Eva, and Vazirani, Vijay V. *Algorithmic game theory*. Cambridge University Press, 2007.

Rüschendorf, L. Solution of a statistical optimization problem by rearrangement methods. *Metrika*, 30(1):55–61, 1983.

Shaked, Moshe and Shanthikumar, J George. Parametric stochastic convexity and concavity of stochastic processes. *Annals of the Institute of Statistical Mathematics*, 42(3):509–531, 1990.

Shioura, Akiyoshi. On the pipage rounding algorithm for submodular function maximization view from discrete convex analysis. *Discrete Mathematics, Algorithms and Applications*, 1(01):1–23, 2009.

Soma, Tasuku, Kakimura, Naonori, Inaba, Kazuhiro, and Kawarabayashi, Ken-ichi. Optimal budget allocation: Theoretical guarantee and efficient algorithm. In *Proceedings of the 31st International Conference on Machine Learning*, pp. 351–359, 2014.

Sviridenko, Maxim. A note on maximizing a submodular set function subject to a knapsack constraint. *Operations Research Letters*, 32(1):41–43, 2004.

Topkis, Donald M. *Supermodularity and complementarity*. Princeton University Press, 1998.

Tzoumas, Vasileios, Amanatidis, Christos, and Markakis, Evangelos. A game-theoretic analysis of a competitive diffusion process over social networks. In *Internet and Network Economics*, pp. 1–14. 2012.

Varian, Hal R. Position auctions. *International Journal of Industrial Organization*, 25(6):1163–1178, 2007.

Vetta, Adrian. Nash equilibria in competitive societies, with applications to facility location, traffic routing and auctions. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 416–425, 2002.

Yuan, Shuai, Wang, Jun, and Zhao, Xiaoxue. Real-time bidding for online advertising: measurement and analysis. In *Proceedings of the 7th International Workshop on Data Mining for Online Advertising*, pp. 3:1–3:8, 2013.

Zhao, Lan and Nagurney, Anna. A network equilibrium framework for internet advertising: Models, qualitative analysis, and algorithms. *European Journal of Operational Research*, 187(2):456–472, 2008.

A. Proofs omitted in the body

Here, we give all missing proofs for the presented propositions.

Proof of Proposition 1. This follows from explicit formula (10) and the fact that (9) is monotone-nondecreasing submodular function with component-wise concavity in x_i (Soma et al., 2014). □

Proof of Proposition 2. We find a potential Φ of the form

$$\Phi(x) = \sum_{t \in T} \sum_{I \subseteq [n]} \alpha(I) \prod_{i \in I} P_i(x_i, t),$$

where $\alpha(I)$ denotes a coefficient to be determined. Condition (11) of potential is invariant under translation; thus, without loss of generality, we may assume $\alpha(\emptyset) = 0$.

For each $i \in [n]$, we have

$$\begin{aligned} &\Phi(x_i', x_{-i}) - \Phi(x_i, x_{-i}) \\ =& \sum_{t \in T} (P_i(x_i, t) - P_i(x_i', t)) \sum_{i \in I \subseteq [n]} \alpha(I) \prod_{j \in I \setminus \{i\}} P_j(x_j, t) \end{aligned}$$

and

$$\begin{aligned} &f_i(x_i', x_{-i}) - f_i(x_i, x_{-i}) \\ =& \sum_{t \in T} (P_i(x_i, t) - P_i(x_i', t)) \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \prod_{j \prec_\sigma i} (1 - P_j(x_j, t)). \end{aligned}$$

To hold equality (11), the coefficient α should satisfy the following relation.

$$\begin{aligned} &\sum_{i \in I \subseteq [n]} \alpha(I) \prod_{j \in I \setminus \{i\}} P_j(x_j, t) \\ &\qquad = \frac{1}{n!} \sum_{\sigma \in \mathcal{S}_n} \prod_{j \prec_\sigma i} (1 - P_j(x_j, t)). \end{aligned}$$

By comparing these terms, for all $I \subseteq V$ that contain $i \in V$, we have[8]

$$\begin{aligned} \alpha(I) &= \frac{(-1)^{|I|+1}}{n!} |\{\sigma \in \mathcal{S}_n : j \prec_\sigma i \ (\forall j \in I \setminus \{i\})\}| \\ &= \frac{(-1)^{|I|+1}}{|I|}. \end{aligned}$$

[8]The number of permutations is evaluated as follows. Let $I = \{1, \dots, |I|\}$. Consider a permutation of symbols $(x, \dots, x, |I| + 1, \dots, n)$. By replacing the last appeared x with i and the other x with distinct $j \in I$, we obtain a permutation that satisfies $j \prec_\sigma i \ (\forall j \in I \setminus \{i\})$. This shows that there is a one-to-one correspondence between $\mathcal{S}_n / \mathcal{S}_I \times \mathcal{S}_{I \setminus \{i\}}$ and $\{\sigma \in \mathcal{S}_n : j \prec_\sigma i \ (\forall j \in I \setminus \{i\})\}$. Therefore, $|\{\sigma \in \mathcal{S}_n : j \prec_\sigma i \ (\forall j \in I \setminus \{i\})\}| = n!(|I| - 1)!/|I|! = n!/|I|$.

It is important that the above-derived $\alpha(I)$ is independent to the choice of i. Therefore

$$\Phi(x) = \sum_{t \in T} \sum_{\emptyset \neq I \subseteq [n]} (-1)^{|I|-1} \frac{1}{|I|} \prod_{i \in I} P_i(x_i, t) \tag{17}$$

satisfies condition (11) for all i. □

Proof of Proposition 4. Let $x^* = (x_1^*, \dots, x_n^*) \in \mathcal{D}$ be the social optimal allocation, and let $x = (x_1, \dots, x_n) \in \mathcal{D}$ be a Nash equilibrium. Let $x^{*(i)} = (x_1^*, \dots, x_i^*, 0, \dots, 0)$ for $i = 0, \dots, n$. Then we have

$$\begin{aligned} F(x^*) - F(x) \le& F(x^* \vee x) - F(x) \\ &= \sum_{i=1}^n F(x^{*(i)} \vee x) - F(x^{*(i-1)} \vee x) \\ &\le \sum_{i=1}^n F(x_i^* \vee x_i, x_{-i}) - F(x_i, x_{-i}) \\ &\le \sum_{i=1}^n F(x_i^* \vee x_i - x_i, x_{-i}) - F(0, x_{-i}). \end{aligned}$$

Here, the first line is obtained by monotonicity, the second line is the telescoping sum, the third line is by submodularity, and the last line is by component-wise concavity. Since $x_i^* \vee x_i - x_i \le x_i^*$ and by the definition of $\mathcal{D}_i$, we have $x_i^* \vee x_i - x_i \in \mathcal{D}_i$.

By condition 3) of utility games, we have

$$F(x_i^* \vee x_i - x_i, x_{-i}) - F(0, x_{-i}) \le f_i(x_i^* \vee x_i - x_i, x_{-i}).$$

Furthermore, since x is a Nash equilibrium, we have

$$f_i(x_i^* \vee x_i - x_i, x_{-i}) \le f_i(x_i, x_{-i}). \tag{18}$$

Therefore we have

$$F(x^*) - F(x) \le \sum_{i=1}^n f_i(x) = F(x).$$

This shows $F(x^*) \le 2F(x)$, i.e., PoA ≤ 2. □

Proof of Proposition 5. By the definition of F, condition 2) is obvious. Here, we prove conditions 1) and 3).

1). Corollary 2.6.3 in Topkis (Topkis, 1998) shows that the product of nonnegative nonincreasing supermodular (i.e., negative of submodular) functions is also a nonnegative nonincreasing supermodular function. Since each $P_i(x_i, t)$ is a nonnegative nondecreasing submodular function, the product

$$1 - \prod_{i=1}^n (1 - P_i(x_i, t)) \tag{19}$$

is a nonnegative nondecreasing submodular function. Furthermore, since $P_i(x_i, t)$ satisfies component-wise concavity and the variables $x_1, \dots, x_n$ appear separately in (19), it also satisfies component-wise concavity. This proves condition 1).

3). Since $f_i(0, x_{-i}) = 0$, we have

$$f_i(x) = f_i(x) - f_i(0, x_{-i}) = \Phi(x) - \Phi(0, x_{-i}). \quad (20)$$

By the form of the potential function (17), we have

$$f_i(x) = \sum_{t \in T} P_i(x_i, t) \sum_{I \subseteq [n] \setminus \{i\}} \frac{(-1)^{|I|}}{|I|+1} \prod_{j \in I} P_j(x_j, t).$$

On the other hand, by the definition (13), we have

$$\begin{aligned} &F(x_i, x_{-i}) - F(0, x_{-i}) \\ &\quad = \sum_{t \in T} P_i(x_i, t) \prod_{j \in [n] \setminus \{i\}} (1 - P_j(x_j, t)). \end{aligned}$$

We prove the inequality between these two formulas. Consider the function

$$\begin{aligned} \Psi_i(x, t; z) &= \prod_{j \in [n] \setminus \{i\}} (1 - z P_j(x_j, t)) \\ &= \sum_{I \subseteq [n] \setminus \{i\}} (-1)^{|I|} z^{|I|} \prod_{j \in I} P_j(x_j, t). \quad (21) \end{aligned}$$

Then we have the following representations.

$$F(x_i, x_{-i}) - F(0, x_{-i}) = \sum_{t \in T} P_i(x_i, t) \Psi_i(x, t; 1), \quad (22)$$

$$f_i(x) = \sum_{t \in T} P_i(x_i, t) \int_0^1 \Psi_i(x, t; z) dz. \quad (23)$$

By definition (21) of Ψ_i, it is a monotone nonincreasing function in $z \in [0, 1]$. In particular, $\Psi(X, t; z)$ takes minimum at $z = 1$. Thus, by comparing (22) and (23), we obtain 3). □

Proof of Proposition 7. Let $x^+ \in \operatorname{argmax}\{\Phi(x) : x \in \mathcal{D}\}$, which is a Nash equilibrium. We estimate the social utility at x^+. Similar to the proof of Proposition 5, we define

$$\Psi(x^+, t; z) = \frac{1}{z} \left(1 - \prod_{i \in [n]} (1 - z P_i(x_i^+, t)) \right). \quad (24)$$

By induction in n, we can prove that $\Psi(x^+, t; z)$ is a monotone nonincreasing function in $z \in [0, 1]$. Thus, since we have

$$F(x^+) = \sum_{t \in T} \Psi(x^+, t; 1), and$$

$$\Phi(x^+) = \sum_{t \in T} \int_0^1 \Psi(x^+, t; z) dz, \quad (25)$$

we have $F(x^+) \le \Phi(x^+)$. We now prove the opposite inequality. To this end, we use

$$\max_{x \in \mathcal{D}} \frac{\Phi(x)}{F(x)} \le \max_{a \in [0,1]^n} \int \Xi(a; z) dz,$$

where

$$\Xi(a; z) = \frac{(1/z)(1 - \prod_{i \in [n]} (1 - z a(i)))}{1 - \prod_{i \in [n]} (1 - a(i))}.$$

By computing $\partial \Xi(a; z) / \partial a(j)$, it follows that $\Xi(a; z)$ is a monotone nondecreasing function in $a \in [0, 1]^n$. Thus we have

$$\begin{aligned} \max_{a \in [0,1]^n} \int_0^1 \Xi(a; z) dz &= \int_0^1 \Xi(1; z) dz \\ &= \int_0^1 \frac{1 - (1 - z)^n}{z} dz = H_n. \end{aligned}$$

Thus we have $\Phi(x) \le H_n F(x)$ for all $x \in \mathcal{D}$. Therefore, for the best utility configuration x^*, we have

$$F(x^*) \le \Phi(x^*) \le \Phi(x^+) \le H_n F(x^+).$$

This shows PoS $\le H_n$. □

Proof of Proposition 9. When the approximate best response dynamics terminates, we have

$$f_i(x_i, x_{-i}) \ge \eta f_i(x_i', x_{-i}) - \epsilon.$$

for all $x_i' \in \mathcal{D}_i$. Therefore the obtained solution is an (η, ϵ)-approximate Nash equilibrium.

We evaluate the number of rounds. Let $x^{(\nu)}$ be the solution at the ν-th round. If the algorithm does not terminate at this round, some advertiser i's utility is increased by at least ϵ. Therefore, since the game is a potential game, we have

$$\Phi(x^{(\nu+1)}) - \Phi(x^{(\nu)}) \ge \epsilon.$$

Therefore, the number of rounds is at most $(\max_x \Phi(x))/\epsilon$. By the proof of Proposition 7, we have $\max_x \Phi(x) \le H_n \max_x F(x) \le |T| H_n$. Thus the approximate best response dynamics converges in at most $|T| H_n / \epsilon$ iterations. □

Proof of Proposition 10. Since (20) holds, we only need to consider how to compute $\Phi(x)$ efficiently. We use the integral representation (25): Since each integrant $\Psi(x, t; z)$ is a polynomial in z, we expand (24) and substitute to (25) to compute $\Phi(x)$ as follows.

We first compute $P_i(x, t)$ for all $i \in [n]$ and $t \in T$; This requires $O(n|E|)$ time. We then expand $\Psi(x, t; z)$ to obtain the representation

$$\Psi(x, t; z) = c_0(x, t) + c_1(x, t) z + \cdots + c_{n-1}(x, t) z^{n-1}. \quad (26)$$

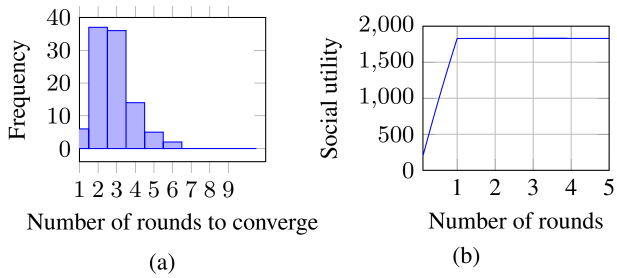

Figure 3. Number of rounds to converge in power-law graph.

Then we can compute $\Phi(x)$ via

$$\Phi(x) = \sum_{t \in T} \left(c_0(x,t) + \frac{c_1(x,t)}{2} + \cdots + \frac{c_{n-1}(x,t)}{n} \right).$$

To obtain the representation (26), we can use a divide-and-conquer method with fast Fourier transform: divide the factors, compute the products recursively, and then compute the product by the fast Fourier transform. The complexity of this method is $O(n(\log n)^2)$ because the following equation

$$g(n) = 2g(n/2) + O(n \log n)$$

has a solution $g(n) = O(n(\log n)^2)$. Therefore, we obtain an $O(|T|n(\log n)^2)$ time algorithm for computing $\Phi(x)$. Note that this divide-and-conquer approach can be also found in Li, Saha, and Deshpande (Li et al., 2009). □

Proof of Proposition 11. We only have to modify (18) in the proof of Proposition 4 as follows.

$$\eta f_i((x_i^* \vee x_i) - x_i, x_{-i}) - \epsilon \leq f_i(x_i, x_{-i}).$$

This proves the proposition. □

B. Additional experimental results

Here, we show some experimental results which are omitted in the main body.

Additional experimental results for Section 8.1 In Section 8.1, we give a typical behavior of the proposed model by using a random bipartite graph whose right vertices has a constant number of edges. Here, we give a result for the same experiment for a random power-law bipartite graph.

A graph is constructed as follows: Let S be the set of left vertices and T be the set of right vertices. We specify degrees of the right vertices $t \in T$ by a power-law distribution. Then, we add random edges which are consistent with specified degrees. This construction is called *configuration model*.

The results are shown in Figure 3. As same as the results shown in Figure 2, most instances are converged in four rounds, and the process almost converges at the first round.

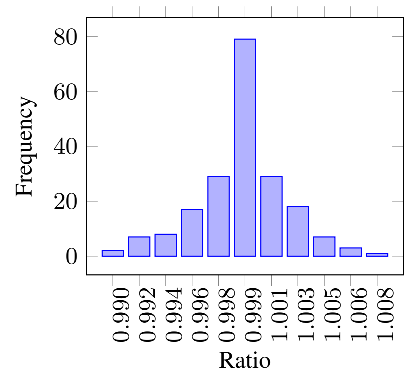

Figure 4. Histogram of the utility by exact minimization / by approximate minimization.

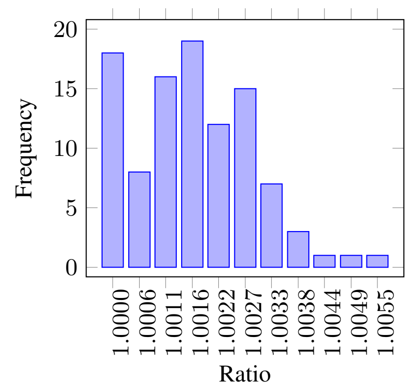

Figure 5. Histogram of the optimal social utility / utility of an obtained approximate Nash equilibrium.

We can observe that the distributions of the number of rounds to converge (Figures 2 (a) and 3 (a)) are slightly different. In a random power-law bipartite graph, the process converges more quickly than that for a random uniform bipartite graph.

Additional experimental results for Section 8.3 In Section 8.3, we described that the personal utilities do not change regardless of the exact maximization or the approximate maximization in each step. Here, we give a raw data for this experiment. Figure 4 is a histogram of this data. The ratio is contained in a small range, $[0.990, 1.008]$, and it is sharply concentrated at a very small range, $[0.998, 1.001]$.

Additional experimental results for Section 8.4 In Section 8.4, we described that the the ratio of the optimal social utility and the utility of an obtained approximate Nash equilibrium is very close. Since the approximate price of anarchy is 4.16, this result is much better than the theory. Here, we give a raw data for this experiment. Figure 5 is a histogram of this data. The ratio is contained in a small range, $[1.000, 1.006]$, and the majority is contained in a more smaller range, $[1.000, 1.003]$.