12-2022

Mining competitively-priced bundle configurations

Ezekiel Ong YOUNG
Singapore Management University, eoyoung.2019@economics.smu.edu.sg

Hady W. LAUW
Singapore Management University, hadywlauw@smu.edu.sg

Mining Competitively-Priced Bundle Configurations

Ezekiel Ong Young
School of Computing and Information Systems
Singapore Management University
eoyoung.2019@smu.edu.sg

Hady W. Lauw
School of Computing and Information Systems
Singapore Management University
hadywlauw@smu.edu.sg

***Abstract*—We examine the bundle configuration problem in the presence of competition. Given a competitor's bundle configuration and pricing, we determine what to bundle together, and at what prices, to maximize the target firm's revenue. We highlight the difficulty in pricing bundles and propose a scalable alternative and an efficient search heuristic to refine the approximate prices. Furthermore, we extend the heuristics proposed by previous work to accommodate the presence of a competitor. We analyze the effectiveness of our proposed models through experimentation on real-life ratings-based preference data.**

I. Problem

A data-driven approach to bundling is prevalent due to the availability of information [5] [4]. However, the current work in computational bundling often assumes the setting of a monopoly. Novelly, our work analyzes it from the perspective of a target firm, facing an existing competitor firm selling the same inventory of items, potentially in different bundle configurations and at different prices. Assuming consumer preferences to be initially known and the firm's objective to maximize revenue, we have the *bundle configuration problem*, wherein the firm has *N* items and must decide which items to bundle together and at what price to sell the bundles.

We introduce several terminologies. Willingness to pay is the maximum amount that a consumer is willing to pay for an item. Let consumer *u*'s willingness to pay for bundle *(A)* be $w_{u,(A)} \in \mathbb{R}_0^+$, in dollars. Consumer surplus, CS, is then the difference between the willingness to pay and the price. We assume consumers are rational and seek to maximize utility.

We consider a set of *M* consumers $U = \{u_1, ..., u_M\}$ and a set of *N* items $I = \{i_1, ..., i_N\}$. *W* is an $M \times N$ matrix, where each element $w_{u,i} \in \mathbb{R}_0^+$ indicates how much a consumer *u* is willing to pay for an item *i*. *W* is presumably known and specified as input. With the willingness to pay for each of the individual items in the matrix *W*, we derive the willingness to pay for any possible bundle as the sum of the willingness to pay for each of the items in the bundle.

Consumer Decision. We consider the impact of a single existing competitor[1] on the target firm's bundle configuration and pricing. We assume that consumers can only buy from either the target firm or the competitor but not both, which is reasonable in markets with considerable switching costs.

Table I provides an illustration of the consumer decision in the presence of competition. Given a set of items {*A*, *B*}, let *(A, B)* denote a bundle containing *A* and *B* together. $P_{(.)}$ denotes the price of a bundle. From the table, we observe that u_1's consumer surplus from buying from the target firm is ($22 - $18) due to *(A, B)* + ($8 - $7) due to *(C)*, totalling $5. Meanwhile her consumer surplus from buying from the competition would have been ($16 - $15) due to *(A)* + ($6 - $5) due to *(B)* + ($8 - $6) due to *(C)*, totalling $4. As a result, u_1 will choose to buy *(A, B)* and *(C)* from the target firm to maximize her consumer surplus. Further observe that u_2's consumer surplus from the target firm is $1 while her consumer surplus from the competitor firm is $6. As a result, u_2 buys from the competition and buys only *(B)* and *(C)*. Thus, a consumer purchases a bundle from the target firm if the consumer's willingness to pay (for that bundle) *exceeds* the price (offered by the target firm) *and* if the target firm's consumer surplus (for that particular consumer) is *greater than* the competition's consumer surplus. When the target and competitor firms' prices are identical, purchases can be apportioned equally.

Consumer	Willingness to Pay				Target Firm		Competitor Firm		
	$w_{u,A}$	$w_{u,B}$	$w_{u,AB}$	$w_{u,C}$	$P_{AB} = 18$	$P_C = 7$	$P_A = 15$	$P_B = 5$	$P_C = 6$
u_1	$16	$6	$22	$8	✓	✓			
u_2	$9	$9	$18	$8				✓	✓

TABLE I: Consumer Decision in the Presence of Competition

***k*-sized Bundle Configuration Problem.** We denote a bundle $b \subseteq I$ to be a set of items. Notation-wise, a size-*k* bundle refers to a bundle *b* of exactly $|b| = k$ items while a *k*-sized bundle refers to a bundle *b* of size $1 \leq |b| \leq k$. Furthermore, let r_b denote the target firm's revenue of bundle *b*, and C_I, P_C denote the bundle configuration and the pricing of the competition's bundles respectively.

Problem 1 (k-SIZED BUNDLING). *Given W, C_I, P_C and an integer $k \geq 1$, find the bundle configuration χ_I and corresponding bundle pricing P_χ, containing k-sized bundles meeting the conditions:*

1) $\cup \chi_I = I$, i.e., the union of sets in χ_I is I
2) $\forall b_1, b_2 \in \chi_I, b_1 \cap b_2 \neq \emptyset$ implies $b_1 = b_2$
3) $\sum_{b \in \chi_I} r_b$ *is maximized, i.e., no other configuration of I will yield a higher overall revenue.*

The first condition enforces how the combination of all the bundles should result in the full set of items. The second condition enforces each item belonging to only one bundle. The last condition imposes the revenue maximization objective. Bundle configuration satisfying these conditions would be regarded *optimal*. The parameter *k* limits the maximum bundle size. Here, we adopt the standard assumptions of bundling [3].

[1]The general approach could be extended to multiple existing competitors with added computations, which would be explored in future work.

Proposed Approach. Our objective to arrive at the optimal bundle configuration and the corresponding prices for the bundles can be decomposed into *two phases* dealing with specific research issues. To determine the optimal bundling configuration, we must first optimally price potential bundles which provides the expected revenue for these bundles. Thereafter, we can devise an algorithmic approach to determining the optimal bundle configuration.

- In Section II-A, we investigate the pricing phase. For a given a bundling configuration χ_I, competition's bundle configuration C_I and bundle pricing P_C, we seek to estimate the prices p_b for each bundle, *b*, to maximize the target firm's revenue.
- In Section II-B, we explore the bundling phase. Given a set of candidate bundles $\{b\}$ defined over *I*, their corresponding estimated revenues $\{r_b\}$, competition bundle configuration C_I and competition's bundle pricing P_C, we seek to find the configuration χ_I that maximizes the target firm's revenue.

II. Methodology

A. Competitive Pricing

To illustrate the difficulty of pricing, we observe the example provided in Section I. To attract u_2, the target firm needs to decrease its prices to increase u_2's consumer surplus from the target firm. Thus, a consumer's total surplus is a function of the prices of the bundles. This induces dependencies between the prices of bundles since the consumer's decision is based off comparing the relative consumer surplus overall.

Hence, pricing bundles independently, as described in [3], would fail to determine the exact revenue generated by that bundle. For example, merely considering the pricing of bundle (A, B) in the earlier example, the expected revenue would be $\$18 \times 2 = \$36$ since both u_1 and u_2 can buy the bundle. However, the actual revenue generated by the bundle is only \$18, as shown in Table I, due to the presence of competition.

We observe two issues that arise. First, the revenue-maximization objective overestimates the prices that consumers would accept now that they have more options. Second, independent pricing is insensitive to the competition pricing.

λ-Surplus Maximization Pricing. We propose the *λ-Surplus Maximization Pricing*, consequently referred to as *λ-pricing*, which utilizes the following optimization to determine the price:

$$\max_p \sum_u x_u \cdot p + \lambda \cdot x_u \cdot [w_u - p]$$
$$\text{s.t. } x_u = 1 \text{ if } p \leq w_u \text{ else } x_u = 0, \forall u$$

where x_u is a dummy variable to denote the *u*-th consumer's choice and w_u represented the *u*-th consumer's willingness to pay.

Intuitively, instead of merely maximizing revenue, we accord a certain weight, λ, to the consumer's surplus, $w_u - p$. When $\lambda = 0$, the *λ-pricing* will reduce to the *revenue-maximizing price*, which is the pricing utilized by [3] to determine the optimal price for a monopoly. For $\lambda > 0$, the objective is no longer to maximize the target firm's revenue per se, as that revenue may not be realizable. Instead, we anticipate a lower expected revenue ($\sum_u x_u \cdot p^*$, where p^* is the *λ-price*). By intentionally factoring a fraction of the consumer surplus when determining the price, we ensure that consumers would gain ample consumer surplus from the target firm thus making them more likely to purchase from the target firm than from the competition. This narrows the gap between the expected and actual revenues, creating a more reliable form of pricing. Experiments in Section III demonstrate the performance of the *λ-price* against the *revenue maximizing price* for different values of λ. Since the *λ-price* is merely a linear transformation in M, the computational complexity of the *λ-price* is the same as the *revenue maximizing price*, which is $O(M)$ [3].

Price Refinement. We further propose a *price refinement* to search for prices around the *λ-price*. This has a two-fold benefit. For one, the refinement provides a pricing closer to the optimal while needing less computational resources as a complete search over all possible prices. Refinement can only be conducted when given a particular bundle configuration due to the dependency between the prices. For another, by paying attention to the competition pricing during refinement, we can ensure competitiveness. Experiments in Section III demonstrate the gains from utilizing the price refinement.

B. Competition-Aware Bundling

While the *2*-sized bundle configuration problem has an optimal solution of polynomial time [3], the *k*-sized bundle configuration problem, where $k \geq 3$, is NP-hard. We adapt the heuristic for the *k*-sized bundle configuration problem from [3] to our research that considers competition as in Algorithm 1.

Each item is modelled as a node in a graph, each bundle of two as an edge between two nodes and each bundle of one as an edge from a node to itself. Given a graph $G(V, E)$, each vertex in V corresponds to an item in *I*, i.e., $|V| = N$. The set of edges E contains $N + N(N-1)/2$ edges since there are N original bundles and $N(N-1)/2$ combination bundles. By weighting each edge by the *λ-pricing* revenue of the bundle, we obtain the maximum weighted matching in G [1].

Note a key difference in the nature of revenue. In [3], the calculated revenue is the actual revenue as the consumers do not have an alternative due to the assumption of a monopoly. In this research, the revenue obtained from the *λ-pricing* only provides the expected revenue. As such, we further propose to conduct the *price refinement* after the maximum weight matching is obtained (see Section II-A).

The diminishing geometric progression of this algorithm justifies the complexity to be still $O(N^{2.5} + MN^2)$ [3].

III. Experiments

Data. We utilize the *Books*, *Electronics* and *Video Games* datasets from the Amazon Review Data provided by [2] to populate our willingness to pay matrix, *W*. In the data, each rating is assigned by a user to a product, on a scale of 1 (lowest) to 5 (highest). Since ratings are extremely sparse,

Algorithm 1 Competition-Aware Bundling Algorithm

```
Initialize χ_I to be the set of size-1 bundles.
Initialize R with the revenue of components.
while true do
    Construct a graph G with χ_I as vertices
    Populate G with edges involving newly-formed bundles.
    Compute the weight of each edge (see Section II-A)
    Obtain the maximum weight matching S in G.
    Update the prices of the chosen bundles in S using the price refinement
    Compute R', the weight or revenue of S.
    if R' ≤ R then Break.
    R ← R'
    for each selected edge in S do
        Remove the edge's vertices from χ_I.
        Collapse the edge into a new vertex in χ_I.
Return χ_I.
```

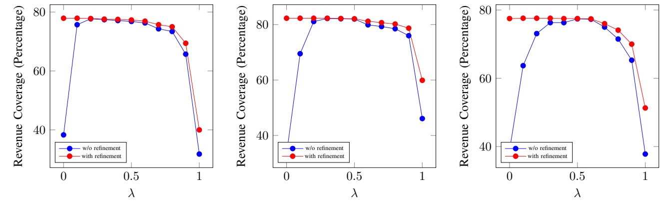
we obtain the top 100 items, in terms of the number of reviews, and the corresponding top 1000 consumers, who have reviewed these items the most. i.e., $N = 100, M = 1000$. We adopt the same function as in [3] for transforming the given ratings into willingness to pay.

Revenue Coverage. A key metric would be the proximity to the absolute maximum. The combined willingness to pay in the input matrix W is the upper bound of revenue. The *revenue coverage* of an algorithm is the ratio (expressed in percentage) of its revenue to the total willingness to pay. The ideal score would be 100%.

Objectives. First, we compare the performance of the various pricing mechanisms, *revenue-maximizing price* against λ-*pricing* (for different values of λ). Second, we demonstrate the utility of using the *price refinement* on λ-*pricing*. Third, we compare the competition-aware bundling algorithm against the monopolistic algorithm proposed in [3]. Finally, we highlight the scalability of our overall algorithm.

Pricing Comparison and Refinement. To simulate competition, we assign the competitor with a random bundle configuration, with bundles priced using the *revenue-maximizing price*. Using the average revenue from 25 randomly generated bundle configurations, Figure 1 shows the comparison of λ-*pricing* for different values of λ, before and after the *price refinement*. We observe that the *revenue-maximizing price* (i.e., $\lambda = 0$, without refinement) performs poorly as compared to the proposed λ-*pricing* (i.e., $\lambda > 0$). Moreover, we observe the benefit of the proposed *price refinement*, which generally increases the average revenue obtained.

Bundle Configuration. Using the average revenue from 10 experiments (corresponding to different bundle configurations of competitor firm), Table II demonstrates the efficacy of the proposed competition-aware Algorithm 1, against the monopolistic algorithm from [3] when considering the presence of a competitor in the market. Evidently, by paying attention to the



(a) *Books* (b) *Electronics* (c) *Video Games*

Fig. 1: Revenue Coverage: λ-pricing without vs. with refinement

	Books	*Electronics*	*Video Games*
Monopolistic Bundling [3]	0%	59.0%	38.1%
Competition-Aware Bundling ($\lambda = 0.1$)	**78.3%**	**82.3%**	**77.9%**

TABLE II: Revenue Coverage: Monopolistic vs. Competition-Aware

competition, we can garner greater revenue than the original algorithm that ignores the competition.

Scalability. To investigate the scalability of the algorithm, we measure the running time against the number of items and users. Figure 2 shows the scalability with respect to different multiples of items and users respectively. Both axes are in log_2 scale. The general trend of linear growth in log-log axis highlights the polynomial time complexity of the algorithm with respect to the number of items and users. We observe a similar trend in growth which highlights the comparable complexity between the algorithms.

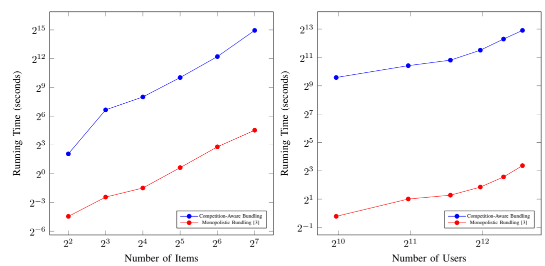


Fig. 2: Scalability of Bundling Algorithms

Acknowledgments

Hady W. Lauw gratefully acknowledges the support by the Lee Kong Chian Fellowship awarded by Singapore Management University.

References

[1] J. Edmonds, "Paths, trees, and Flowers," Canadian Journal of Mathematics, vol. 17, pp. 449–467, 1965.

[2] J. Ni, J. Li, and J. McAuley, "Justifying recommendations using distantly-labeled reviews and fine-grained aspects," Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2019.

[3] L. Do, H. W. Lauw, and K. Wang, "Mining revenue-maximizing bundling configuration," Proceedings of the VLDB Endowment, vol. 8, no. 5, pp. 593–604, 2015.

[4] M. Ettl, P. Harsha, A. Papush, and G. Perakis, "A data-driven approach to personalized bundle pricing and recommendation," Manufacturing amp; Service Operations Management, vol. 22, no. 3, pp. 461–480, 2020.

[5] T. Zhu, P. Harrington, J. Li, and L. Tang, "Bundle recommendation in ecommerce," Proceedings of the 37th international ACM SIGIR conference on Research amp; development in information retrieval, 2014.