An Economics-Based Analysis of RANKING for Online Bipartite Matching

Alon Eden[1], Michal Feldman[2], Amos Fiat[2], and Kineret Segal[2]

[1]Harvard University
[2]Tel Aviv University

Abstract

In their seminal paper, Karp, Vazirani and Vazirani (STOC'90) introduce the online bipartite matching problem, and the RANKING algorithm, which admits a tight $1 - \frac{1}{e}$ competitive ratio. Since its publication, the problem has received considerable attention, including a sequence of simplified proofs. In this paper we present a new proof that gives an economic interpretation of the RANKING algorithm — further simplifying the proof and avoiding arguments such as duality. The new proof gives a new perspective on previous proofs.

1 The Online Bipartite Matching Problem

Consider a bipartite graph $G = (L, R; E)$, where L and R are the sets of left and right vertices, respectively, and $E \subseteq L \times R$ is the set of edges. A matching in G is a subset of the $M \subseteq E$ such that no vertex is incident to more than one edge in M.

In online bipartite matching problem, introduced in the seminal work of Karp, Vazirani and Vazirani [1], the set of edges, E, is initially unknown. The vertices in R are all present at the start, but vertices from L, along with their edges, appear over time. The i'th vertex, $\ell_i \in L$, along with all edges from ℓ_i to vertices in R appear (simultaneously) at the ith time step. Immediately and irrevocably, the online algorithm must decide which, if any, vertex $r \in R$ is to be matched to the new vertex $\ell_i \in L$.

The objective of the online algorithm is to maximize the cardinality of the final matching (after all vertices in L have appeared).

The online bipartite matching problem and variants thereof have received a surge of interest in recent years due to their connection to allocation problems in economic settings, such as internet advertising applications. The new problems are theoretically appealing and practically relevant. See Mehta [2] and references therein for a survey.

For ease of exposition, we assume that $L = \{\ell_1, \ldots, \ell_n\}$, $R = \{\ell_1, \ldots, \ell_n\}$, and G admits a matching of size n. However, the results and proofs in this paper hold for arbitrary bipartite graphs, where $|L|$ and $|R|$ may be arbitrary and n is the size of the maximum matching (n is unknown apriori).

A family of simple greedy algorithms for the online bipartite matching problem match every arriving vertex with an arbitrary unmatched neighbor, if available. (This is a family of algorithms as the arbitrary choices may be different). Every such greedy algorithm outputs a maximal matching, hence has cardinality of at least $n/2$. It is easy to see that this bound is tight; *i.e.*, there exist graphs for which this greedy algorithm cannot achieve more than half of the maximum matching (for example, the first $n/2$ vertices in L are connected to all vertices in R, and the remaining $n/2$ vertices in L are connected to the $n/2$ vertices in R that were matched to the first half of the vertices in L.) A randomized version of the greedy algorithm, which chooses a currently unmatched neighbor uniformly at random (if one exists), cannot improve upon this ratio either (up to lower order terms) [1].

The Randomized RANKING Algorithm. The RANKING algorithm, introduced by Karp, Vazirani and Vazirani [1], is a simple randomized algorithm, which works as follows: it first chooses a random permutation π over the vertices in R. Upon the arrival of a vertex ℓ_i, RANKING matches ℓ_i to the highest-ranked (with respect to π) currently unmatched neighbor of ℓ_i.

Karp et al. proved that RANKING matches at least $(1 - \frac{1}{e})n$ edges in expectation. They also showed that this bound is tight (up to low order terms[1]).

The analysis in the original paper was quite complicated (and imprecise in places). Subsequent papers by Goel and Mehta [4], Birnbaum and Mathieu [5] and Devanur, Jain and Kleinberg [6] simplified the analysis considerably. In this work we present an arguably even simpler proof, which is based on an economic interpretation of the online bipartite matching problem. It bears similarities to the proof of [6], but does not make an explicit use of linear programming duality.

2 An Economics-Based Analysis of RANKING

In what follows we give an economic interpretation of the online bipartite matching. Given a graph $G = (L, R; E)$, vertices of R represent items, and vertices of L represent utility maximizing *unit-demand* buyers.

For every vertex $x \in L \cup R$, let $N(x)$ denote the neighbors of x, e.g.,, $N(\ell_i) = \{r_j \mid (\ell_i, r_j) \in E\}$. Let $v_i(r_j)$ denote the value of buyer ℓ_i for an item r_j. For our purposes assume that $v_i(r_j) = 1$ if

[1]That is, no online algorithm matches more than $(1 - 1/e)n + O(1)$ edges in expectation. Specifically, Feige [3] shows that no online integral matching algorithm matches more than $(1 - \frac{1}{e})n + 1 - \frac{2}{e} + O(\frac{1}{n!})$ edges in expectation.

$r_j \in N(\ell_i)$, and $v_i(r_j) = 0$ otherwise. For a set of items $X \subset R$, let $v_i(X) = \max_{\{r \in X\}} v_i(r)$.

Consider the following process, hereafter referred to as the *market process*: Before the arrival of any buyer, every item r_j is assigned a price p_j. The *utility* a buyer ℓ_i derives from an item r_j is

$$u_i(r_j) = v_i(r_j) - p_j.$$

Buyers arrive online, in arbitrary order. Upon arrival, a buyer chooses an item that maximizes her utility amongst all remaining items (possibly choosing no item). This process induces a matching M, where $(\ell_i, r_j) \in M$ if buyer ℓ_i chooses item r_j. The *social welfare* of a matching M is the sum of buyer valuations for their items, i.e.,

$$SW(M) = \sum_{i \in [n]} v_i(M(i)),$$

where $M(i)$ denotes the item chosen by buyer ℓ_i (possibly an empty set). Since every buyer that receives an item has value 1, the social welfare of a matching M equals the cardinality of M.

The following claim shows a connection between the market process and the randomized RANKING algorithm.

Claim 2.1. *Consider the market process above. Let D be some arbitrary distribution over $[a, b]$, $0 \leq a < b \leq 1$, with no point mass, and choose the price of item r_j, p_j, independently from D, $p_j \sim D$. Then, the resulting distribution over matchings is identical to the one obtained by the randomized RANKING algorithm.*

Proof. Recall that RANKING chooses a random permutation π over R and upon the arrival of a vertex $\ell_i \in L$, ℓ_i is matched to the highest-ranked (according to π) available vertex in R.

Setting a random price, p_j, for items r_j implies that the utility of item r_j is also random $(1 - p_j)$ — given that there is an edge from the buyer to the item. Buyers always choose the maximal utility item. Ergo, a random permutation π over R is equivalent to random prices, chosen from D, to items $r_j \in R$.

The range $[a, b]$, $b \leq 1$, is to avoid negative utility and no point mass implies no equality in pricing or utility. □

We now use Claim 2.1 to prove that RANKING matches at least $(1 - \frac{1}{e})n$ edges in expectation. It suffices to show some distribution D, and prices independently chosen from D, so that the expected social welfare in the market process is at least $(1 - \frac{1}{e})n$.

A particular distribution of interest is the distribution $\widehat{D}$ from which one samples as follows: choose a value w uniformly at random in the interval $[0, 1]$ and return e^{w-1}.

Theorem 2.2. *The market process where prices are sampled independently from $\widehat{D}$, $p_j \sim \widehat{D}$, for $j = 1, \ldots, n$, gives an expected social welfare of at least $(1 - \frac{1}{e})n$.*

Proof. To prove the theorem, we decompose the social welfare into the sum of buyer utilities and seller's revenue (this technique has proved useful in other settings [7, 8, 9]). For every item r_j, let rev_j denote the revenue obtained by r_j (i.e., p_j if the item was purchased and 0 otherwise). For every buyer ℓ_i, let util_i denote the utility of buyer ℓ_i; i.e.,

$$\text{util}_i = \begin{cases} 1 - p_j, & \text{if buyer } \ell_i \text{ purchases item } r_j \in N(\ell_i) \\ 0, & \text{if buyer } \ell_i \text{ does not purchase any item.} \end{cases}$$

Fix some (arbitrary) arrival order of the buyers and a price vector $\mathbf{p} = (p_1, \ldots, p_n)$, and let M be the matching resulting from this process. The following equation gives the resulting social welfare as the sum of the buyer utilities and the total revenue:

$$\sum_{\ell_i \in L} \text{util}_i + \sum_{r_j \in R} \text{rev}_j = \sum_{(\ell_i, r_j) \in M} (1 - p_j) + p_j = |M|. \tag{1}$$

We now give the key lemma used in the proof of Theorem 2.2.

Lemma 2.3. *Assume prices are sampled independently from $\widehat{D}$, $p_j \sim \widehat{D}$. Then, for every edge $(\ell_i, r_j) \in E$,*

$$\mathop{\mathbb{E}}_{\mathbf{w}} [\text{util}_i + \text{rev}_j] \geq 1 - \frac{1}{e}.$$

Before proving Lemma 2.3, we show that it implies Theorem 2.2. Fix a maximum matching M^* and let M be the matching resulting from the market process above. Using Equation (1), linearity of expectation, and Claim 2.3 we get that:

$$\begin{aligned} \mathop{\mathbb{E}}_{\mathbf{w}} [|M|] &= \mathop{\mathbb{E}}_{\mathbf{w}} \left[\sum_i \text{util}_i + \sum_j \text{rev}_j \right] \geq \mathop{\mathbb{E}}_{\mathbf{w}} \left[\sum_{(\ell_i, r_j) \in M^*} (\text{util}_i + \text{rev}_j) \right] \\ &= \sum_{(\ell_i, r_j) \in M^*} \mathop{\mathbb{E}}_{\mathbf{w}} [\text{util}_i + \text{rev}_j] \geq \left(1 - \frac{1}{e}\right) |M^*|. \end{aligned}$$

We conclude the proof of Theorem 2.2 by proving Lemma 2.3.

Proof of Lemma 2.3. Fix some arbitrary order of buyer arrival σ, buyer ℓ_i and item r_j such that $(\ell_i, r_j) \in E$. Sample item prices from $\widehat{D}$. Let M_{-j} denote the matching produced in a market without item r_j (under the same arrival order σ). Let $p = e^{y-1}$ be the price of the item matched to buyer ℓ_i in M_{-j}; if ℓ_i is left unmatched, set $p = 1$ (in this case $p = e^{y-1}$ for $y = 1$). We make the following two simple observations, essentially trivial when thinking in terms of buyers and sellers:

- **Observation (1):** If $p_j < p$, then item r_j is sold. This follows since buyer ℓ_i derives higher utility from item r_j than from its match in M_{-j}, so either item r_j is purchased by an earlier buyer, or buyer ℓ_i buys it upon arrival (note that in the case where $p = 1$ and r_j is unsold upon ℓ_i's arrival, buyer ℓ_i gains by purchasing item r_j for a price $p_j < 1$).

- **Observation (2):** It holds that $\text{util}_i \geq 1-p$. This observation follows since $1-p$ is the utility of buyer ℓ_i in M_{-j}, and the introduction of an additional item into the market (item r_j in our case) can never decrease the utility of any buyer. This last claim holds since, by induction, every buyer faces the same set of items available to her plus, possibly, one additional item. This is obviously true for the first incoming buyer, and remains true subsequently since the introduction of an additional item never induces a buyer to purchase an item previously waived.

Let $y \in [0,1]$ be the value such that $p = e^{y-1}$. It follows that

$$\mathop{\mathbb{E}}_{\mathbf{w}}[\text{rev}_j] = \mathop{\mathbb{E}}_{\mathbf{w}}[p_j \cdot \mathbb{1}[r_j \text{ is sold}]] \geq \mathop{\mathbb{E}}_{\mathbf{w}}[p_j \cdot \mathbb{1}[p_j < p]] = \int_0^y e^{x-1}dx = e^{y-1} - \frac{1}{e} = p - \frac{1}{e},$$

where the inequality follows by observation (1). It now follows from observation (2) that

$$\mathop{\mathbb{E}}_{\mathbf{w}}[\text{util}_i + \text{rev}_j] \geq 1 - p + p - \frac{1}{e} = 1 - \frac{1}{e},$$

as desired.

□

This concludes the proof of Theorem 2.2. □

Note that our proof is closely related to the proof presented by Devanur et al. [6] (and further simplifications by Mathieu [10]), which is based on linear programming duality. Indeed, the utility and revenue terms used in our analysis are essentially scaled versions of the dual variables in [6]. However, the economic interpretation introduced in this paper simplifies the proof even further. In particular, it does not make an explicit use of linear programming duality, and thus eliminates the need to argue about the dual program and its feasibility altogether. Some of the other arguments in [6] are more readily apparent when viewed from the economic perspective.

Acknowledgements. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 866132), and by the Israel Science Foundation (grant number 317/17).

References

[1] Richard M. Karp, Umesh V. Vazirani, and Vijay V. Vazirani. An optimal algorithm for on-line bipartite matching. In *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, May 13-17, 1990, Baltimore, Maryland, USA*, pages 352–358, 1990.

[2] Aranyak Mehta. Online matching and ad allocation. *Foundations and Trends in Theoretical Computer Science*, 8(4):265–368, 2013.

[3] Uriel Feige. Online matching and ad allocation. *Building Bridges II: Mathematics of Laszlo Lovasz. Editors: Barany, Imre, Katona, Gyula O. H., Sali, Attila*, pages 235–255, 2019.

[4] Gagan Goel and Aranyak Mehta. Online budgeted matching in random input models with applications to adwords. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2008, San Francisco, California, USA, January 20-22, 2008*, pages 982–991, 2008.

[5] Benjamin E. Birnbaum and Claire Mathieu. On-line bipartite matching made simple. *SIGACT News*, 39(1):80–87, 2008.

[6] Nikhil R. Devanur, Kamal Jain, and Robert D. Kleinberg. Randomized primal-dual analysis of RANKING for online bipartite matching. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013, New Orleans, Louisiana, USA, January 6-8, 2013*, pages 101–107, 2013.

[7] Michal Feldman, Nick Gravin, and Brendan Lucier. Combinatorial auctions via posted prices. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 123–135, 2015.

[8] Paul Duetting, Michal Feldman, Thomas Kesselheim, and Brendan Lucier. Prophet inequalities made easy: Stochastic optimization by pricing non-stochastic inputs. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 540–551, 2017.

[9] Soheil Ehsani, MohammadTaghi Hajiaghayi, Thomas Kesselheim, and Sahil Singla. Prophet secretary for combinatorial auctions and matroids. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 700–714, 2018.

[10] Claire Mathieu. A css professor blog: A primal-dual analysis of the ranking algorithm. *http://teachingintrotocs.blogspot.co.il/2011/06/primal-dual-analysis-of-ranking.html*, 2011.