Online bipartite matching with imperfect advice

DAVIN CHOO* and THEMIS GOULEAKIS*, National University of Singapore, Singapore
CHUN KAI LING*, Columbia University, USA
ARNAB BHATTACHARYYA, National University of Singapore, Singapore

We study the problem of online unweighted bipartite matching with n offline vertices and n online vertices where one wishes to be competitive against the optimal offline algorithm. While the classic RANKING algorithm of [3] provably attains competitive ratio of $1 - 1/e > 1/2$, we show that no learning-augmented method can be both 1-consistent and strictly better than 1/2-robust under the adversarial arrival model. Meanwhile, under the random arrival model, we show how one can utilize methods from distribution testing to design an algorithm that takes in external advice about the online vertices and provably achieves competitive ratio interpolating between any ratio attainable by advice-free methods and the optimal ratio of 1, depending on the advice quality.

CCS Concepts: • **Theory of computation → Online algorithms**; **Graph algorithms analysis**; **Algorithm design techniques**.

Additional Key Words and Phrases: Learning-augmented algorithms, Online bipartite matching, Random arrival

ACM Reference Format:
Davin Choo, Themis Gouleakis, Chun Kai Ling, and Arnab Bhattacharyya. 2024. Online bipartite matching with imperfect advice. *J. ACM* 37, 4, Article 111 (April 2024), 3 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Finding matchings in bipartite graphs is a mainstay of algorithms research. The area's mathematical richness is complemented by a vast array of applications — any two-sided market (e.g., kidney exchange, ridesharing) yields a matching problem. In particular, the *online* variant enjoys much attention due to its application in internet advertising. Consider a website with a number of pages and ad slots (videos, images, etc.). Advertisers specify ahead of time the pages and slots they like their ads to appear in, as well as the target user. The website is paid based on the number of ads appropriately fulfilled. Crucially, ads slots are available only when traffic occurs on the website and are not known in advance. Thus, the website is faced with the *online* decision problem of matching advertisements to open ad slots.

The classic online unweighted bipartite matching problem by Karp et al. [3] features n offline vertices U and n online vertices V. Each online vertex $v \in V$ reveals its incident edges sequentially upon arrival. With each arrival, one makes an irrevocable decision whether (and how) to match v with a neighboring vertex in U. The final offline graph $\mathcal{G} = (U \cup V, E)$ is assumed to have a

*Equal contribution.

Authors' Contact Information: Davin Choo, davin@u.nus.edu; Themis Gouleakis, tgoule@nus.edu.sg, National University of Singapore, Singapore; Chun Kai Ling, chunkail@cs.cmu.edu, Columbia University, New York, New York, USA; Arnab Bhattacharyya, arnabb@nus.edu.sg, National University of Singapore, Singapore.

ACM 1557-735X/2024/4-ART111
https://doi.org/XXXXXXX.XXXXXXX

largest possible matching of size $n^* \leq n$, and we seek online algorithms producing matchings of size as close to n^* as possible. The performance of a (randomized) algorithm $\mathcal{A}$ is measured by its *competitive ratio*:

$$\min_{\mathcal{G}=(U\cup V,E)} \min_{V\text{'s arrival seq.}} \frac{\mathbb{E}[\#\text{ matches by } \mathcal{A}]}{n^*}, \tag{1}$$

where the randomness is over any random decisions made by $\mathcal{A}$. Traditionally, one assumes the *adversarial arrival model*, i.e., an adversary controls both the final graph $\mathcal{G}$ and the arrival sequence of online vertices.

Since any maximal matching has size at least $n/2$, a greedy algorithm trivially attains a competitive ratio of $1/2$. Indeed, Karp et al. [3] show that no deterministic algorithm can guarantee better than $1/2 - o(1)$. Meanwhile, the randomized RANKING algorithm of Karp et al. [3] attains an asymptotic competitive ratio of $1 - 1/e$ which is also known to be optimal [1–4].

In practice, *advice* (also called predictions or side information) is often available for these online instances. For example, online advertisers often aggregate past traffic data to estimate the *future* traffic and corresponding user demographic. While such advice may be imperfect, it may nonetheless be useful in increasing revenue and improving upon aforementioned worst-case guarantees. Designing algorithms that utilize such advice in a principled manner falls under the research paradigm of *learning-augmented algorithms*. A learning-augmented algorithm is said to be (i) *a-consistent* if it is *a-competitive* with perfect advice and (ii) *b-robust* if it is *b-competitive* with arbitrary advice quality.

Goal 1. *Let β be the best-known competitive ratio attainable by any classical advice-free online algorithm. Can we design a learning-augmented algorithm for the online bipartite matching problem that is 1-consistent and β-robust?*

Clearly, Goal 1 depends on the form of advice as well as a suitable measure of its quality. Setting these technicalities aside for now, we remark that Goal 1 strikes the best of all worlds: it requires that a perfect matching be obtained when the advice is perfect, while not sacrificing performance with respect to advice-free algorithms when faced with low-quality advice. In other words, there is potential to benefit, but no possible harm when employing such an algorithm.

2 RESULTS

1. Impossibility under adversarial arrivals

We show that under adversarial arrivals, learning augmented algorithms, no matter what form the advice takes, cannot be both 1-consistent and strictly more than 1/2-robust. The latter is worse than the competitive ratio of $1 - 1/e$ guaranteed by known advice-free algorithms [3].

2. Achieving Goal 1 under the random arrival model

We propose an algorithm TestAndMatch achieving Goal 1 under the weaker *random arrival model*, in which an adversary controls the online vertices V but its arrival order is randomized. Our advice is a *histogram over types* of online vertices; in the context of online advertising this corresponds to a forecast of the user demographic and which ads they can be matched to. TestAndMatch assumes perfect advice while simultaneously testing for its accuracy via the initial arrivals. If the advice is deemed useful, we mimic the matching suggested by it; else, we revert to an advice-free method. The testing phase is kept short (sublinear in n) by utilizing state-of-the-art L_1 estimators from distribution testing. We analyze our algorithm's performance as a function of the quality of advice, showing that its competitive ratio gracefully degrades to β as quality of advice decays. To the best of our knowledge, our work is the first that shows how one can leverage techniques from the property testing literature to designing learning-augmented algorithms.

REFERENCES

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

[2] Gagan Goel and Aranyak Mehta. 2008. Online budgeted matching in random input models with applications to Adwords. In *SODA*, Vol. 8. 982–991.

[3] Richard M Karp, Umesh V Vazirani, and Vijay V Vazirani. 1990. An optimal algorithm for on-line bipartite matching. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing*. 352–358.

[4] Vijay V Vazirani. 2022. Online Bipartite Matching and Adwords (Invited Talk). In *47th International Symposium on Mathematical Foundations of Computer Science (MFCS 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

Received 26 April 2024