ARTICLE IN PRESS

YINCO:4338

Information and Computation ••• (••••) •••–•••

Contents lists available at ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

Improved bounds for randomized preemptive online matching

Leah Epstein [a], Asaf Levin [b], Danny Segev [c,*], Oren Weimann [d]

[a] Department of Mathematics, University of Haifa, 31905 Haifa, Israel
[b] Faculty of Industrial Engineering and Management, The Technion, 32000 Haifa, Israel
[c] Department of Statistics, University of Haifa, 31905 Haifa, Israel
[d] Department of Computer Science, University of Haifa, 31905 Haifa, Israel

ARTICLE INFO

Article history:
Received 25 January 2014
Received in revised form 3 May 2017
Available online xxxx

Keywords:
Online matching
Competitive analysis
Lower bounds

ABSTRACT

Preemptive online algorithms for the *maximum matching* problem maintain a valid matching M while edges of the underlying graph are presented one after the other. When presented with an edge e, the algorithm should decide whether to augment the matching M by adding e (in which case e may be removed later on) or to keep M in its current form without adding e (in which case e is lost for good). The objective is to eventually hold a matching M with maximum weight.
The main contribution of this paper is to establish new lower and upper bounds on the competitive ratio achievable by randomized preemptive online algorithms:

- We provide a lower bound of $1 + \ln 2 \approx 1.693$ on the competitive ratio of any randomized algorithm for the maximum cardinality matching problem.
- We devise a randomized algorithm that achieves an expected competitive ratio of 5.356 for maximum weight matching.

1. Introduction

In the *maximum matching* problem, we are given an undirected graph $G = (V, E)$ whose edges are associated with non-negative weights. A set of edges $M \subseteq E$ is called a *matching* when no two of them share a common vertex. The objective is to compute a matching of maximum total weight, where the total weight of the output matching is also called the profit of the algorithm. Due to its wide real-life applicability, as well as to its appealing theoretical nature, this computational problem has received a great deal of attention from various communities such as computer science, mathematics, operations research, and economics (see Schrijver's book [1] and references therein for a comprehensive overview of classic work).

As can only be expected, the algorithmic research revolving around maximum matching has expanded from studying the traditional (offline) setting to additional models. In particular, the online setting has extensively been studied over the last few decades [2–8]. In our model, the edges (along with their weights) are presented one by one to the algorithm, which is required to keep a valid matching M at all times. In other words, once an edge e is presented, the algorithm must decide whether to add it to M or not. However, e may be added only if the resulting set of edges $M \cup \{e\}$ remains a valid matching. If the algorithm decides to add the edge e to M, we say that e is *accepted*, and if the algorithm decides not to add e, we

* Corresponding author.
E-mail addresses: lea@math.haifa.ac.il (L. Epstein), levinas@ie.technion.ac.il (A. Levin), segevd@stat.haifa.ac.il (D. Segev), oren@cs.haifa.ac.il (O. Weimann).

https://doi.org/10.1016/j.ic.2017.12.002
0890-5401/

Please cite this article in press as: L. Epstein et al., Improved bounds for randomized preemptive online matching, Inf. Comput. (2017), https://doi.org/10.1016/j.ic.2017.12.002

say that e is *rejected*. With this regard, to formally specify the online setting being considered, the remaining question is whether the acceptance of an edge is permanent or not.

In the non-preemptive model, the decision of whether or not to add any given edge to M is irrevocable, i.e., once an edge is added to the set M of previously accepted edges it can never be removed. The final matching thus consists of all edges that were ever accepted. Alas, in this model, simple examples demonstrate that the competitive ratio of any (deterministic or randomized) algorithm exceeds any function of the number of vertices, meaning that no competitive algorithm exists (see, for example, [3]). That being said, in the unweighted case (where all edge weights are equal, which is also called the *maximum cardinality matching* problem), a greedy approach that accepts an edge whenever possible has a competitive ratio of 2. For deterministic algorithms, this ratio is actually best possible, as shown by Karp, Vazirani, and Vazirani [6].

In the preemptive model, the algorithm is given more freedom by allowing it to remove previously accepted edges from the current matching at any point in time; this event is called *preemption*. Nevertheless, an edge that was either rejected or preempted cannot be re-inserted to the matching later on. As opposed to the non-preemptive model, with this extra freedom competitive algorithms do exist. Specifically, a deterministic algorithm that was proposed by Feigenbaum et al. [4] attains a competitive ratio of 6. Later on, McGregor [8] improved on this finding, by tweaking it into achieving a ratio of $3+2\sqrt{2}\approx 5.828$. On the other hand, Epstein et al. [3] established a lower bound of 4.967 for any deterministic algorithm, which has subsequently been improved by Varadaraja [9] to (a tight bound of) $3+2\sqrt{2}\approx 5.828$.

The upper bound of McGregor and the lower bound of Varadaraja establish a tight bound of $3+2\sqrt{2}\approx 5.828$ on the competitive ratio of any *deterministic* algorithm in the preemptive model. For *randomized* algorithms, the currently best lower bound of $e/(e-1)\approx 1.581$ can be inferred from the work of Karp et al. [6] on a related model, which is further discussed below. Interestingly, their lower bound proof actually works for the unweighted case.

Semi-streaming matching and other related work What seems to have ignited renewed interest in the preemptive online model is the investigation of maximum matching in the *semi-streaming* model, which was introduced by Muthukrishnan [10]. On the other hand, online algorithms are a long-standing and classic research field (including preemptive variants such as those of [11–15]), and the interest in such models is not restricted to its semi-streaming applications. In semi-streaming models, an algorithm performs one or more passes of reading the input as a stream. This means that, in a single pass, it cannot go back and re-examine parts of the input that appear earlier in the stream. The algorithm is allowed, however, to keep a limited amount of information in memory which is based on the input seen thus far. Next, we discuss a more precise definition which is relevant to matching. In the single-pass version of maximum matching problems in the *semi-streaming* model, which is more relevant for our purposes, the edges (along with their weights) are presented one by one to the algorithm, which is allowed to use only $O(n\cdot \mathrm{polylog}(n))$ space to store information at all times (including a potential output, if it wishes to spend some of its limited memory on that), but is not required to hold a valid matching.[1] For matching problems, the stored information is typically just a set of edges, and the possibility to keep in memory *any* small set of edges (rather than only a matching) is what gives this model its added strength in comparison to the preemptive online model. When multiple passes are allowed, the entire sequence of edges is presented multiple times to the algorithm, allowing it to examine the input again, possibly modifying the set of edges stored in memory, and possibly resulting in an improved output. Below, we discuss the case of a single pass, and refer the interested reader to [16] for recent results and a discussion on multiple-pass matching algorithms (see also McGregor [8] and Ahn and Guha [17,18]).

Epstein et al. [3] observed that the semi-streaming algorithms of Feigenbaum et al. [4] and McGregor [8] (of approximation ratios 6 and 5.828, respectively) can actually be viewed as preemptive online algorithms in disguise, and this is what suggested a connection between the two variants. However, the semi-streaming model is not as strict as the preemptive online model, and the subsequent algorithms of Zelke [19] and that of Epstein et al. [3] (whose approximation ratios are 5.585 and 4.91, respectively) are *not* deterministic preemptive online algorithms. Specifically, the latter may simultaneously hold $\Omega(\log n)$ matchings in memory, arguing that their union contains a good matching, while the former keeps several additional edges for each edge in the current matching. It is worth pointing out that the lower bound of Varadaraja [9] on the competitive ratio of deterministic one-pass semi-streaming algorithms also shows that those algorithms cannot be converted into deterministic preemptive online algorithms while maintaining the same approximation ratio. The above-mentioned approximation ratios were improved to $4+\varepsilon$ by Crouch and Stubbs [20], and later on to $3.5+\varepsilon$ by Grigorescu and Monemizadeh, and Zhou [21], using the same algorithm. All those algorithms use a method called *bucketing* in different ways (which we describe below as our randomized algorithm uses this method), leading to relatively high approximation ratios. Finally, Paz and Schwartzman [22], designed a $(2+\varepsilon)$-approximation via a clever implementation of the local ratio technique. It is unclear whether it is possible to convert these algorithms for the semi-streaming model into randomized preemptive online algorithms.

Two variants for online matching We proceed by discussing the differences and similarities between our edge-arrival model and the one-sided vertex-arrival model of Karp et al. [6], which is very relevant for our purposes here. In our model, edges of a general graph are presented one by one. In the one-sided vertex-arrival model, the underlying graph is bipartite, where

[1] Here, n stands for the number of vertices in the underlying graph, meaning in particular that, when the latter is sufficiently dense, most edges cannot be kept in memory.

Please cite this article in press as: L. Epstein et al., Improved bounds for randomized preemptive online matching, Inf. Comput. (2017), https://doi.org/10.1016/j.ic.2017.12.002