Online Matching and Ad Allocation

Online Matching and Ad Allocation

Aranyak Mehta

Google Research
Mountain View, CA 94043
USA
aranyak@google.com

the essence of knowledge

Boston – Delft

Foundations and Trends® in Theoretical Computer Science

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is A. Mehta, Online Matching and Ad Allocation, Foundations and Trends® in Theoretical Computer Science, vol 8, no 4, pp 265–368, 2012

ISBN: 978-1-60198-718-1

Foundations and Trends® in Theoretical Computer Science

Volume 8 Issue 4, 2012

Editorial Scope

Foundations and Trends® in Theoretical Computer Science publishes survey and tutorial articles in the following topics:

- Algorithmic game theory
- Computational algebra
- Computational aspects of combinatorics and graph theory
- Computational aspects of communication
- Computational biology
- Computational complexity
- Computational geometry
- Computational learning
- Computational Models and Complexity
- Computational Number Theory
- Cryptography and information security
- Data structures
- Database theory
- Design and analysis of algorithms
- Distributed computing
- Information retrieval
- Operations Research
- Parallel algorithms
- Quantum Computation
- Randomness in Computation

Information for Librarians

Foundations and Trends® in Theoretical Computer Science, 2012, Volume 8, 4 issues. ISSN paper version 1551-305X. ISSN online version 1551-3068. Also available as a combined paper and online subscription.

Foundations and Trends® in
Theoretical Computer Science
Vol. 8, No. 4 (2012) 265–368

DOI: 10.1561/0400000057

now
the essence of knowledge

Online Matching and Ad Allocation

Aranyak Mehta

Google Research, 1600 Amphitheatre Pkwy, Mountain View, CA 94043, USA, aranyak@google.com

Abstract

Matching is a classic problem with a rich history and a significant impact, both on the theory of algorithms and in practice. Recently there has been a surge of interest in the online version of matching and its generalizations, due to the important new application domain of Internet advertising. The theory of online matching and allocation has played a critical role in designing algorithms for ad allocation. This monograph surveys the key problems, models and algorithms from online matchings, as well as their implication in the practice of ad allocation. The goal is to provide a classification of the problems in this area, an introduction into the techniques used, a glimpse into the practical impact, and to provide direction in terms of open questions. Matching continues to find core applications in diverse domains, and the advent of massive online and streaming data emphasizes the future applicability of the algorithms and techniques surveyed here.

Contents

1

Introduction

A matching in a graph $G(V, E)$ is a set of edges $M \subseteq E$ such that for every $v \in V$, there is at most one edge in M incident on v. A maximum matching is a matching with the largest size. The problem of finding a maximum matching in a graph is a classic one, rich in history and central to algorithms and complexity. The elegance and complexity of the theory of matching is equally complemented by a rich set of important applications; indeed this problem arises whenever we need to connect any pairs of entities, for example, applicants to jobs, spouses to each other, goods to buyers, or organ donors to recipients.

In this monograph we will focus on the online version of the problem, in bipartite graphs. There has been considerable interest recently in online bipartite matching and its generalizations, driven by the important new applications of Ad Allocation in Internet Advertising, corresponding to matching ad impressions to ad slots. We will describe this motivating application first, before giving a brief overview of the history and foundations of matching.

1.1 Ad Allocation

Internet advertising constitutes perhaps the largest matching problem in the world, both in terms of dollars and number of items. Ads are sold either by auction or through contracts, and the resulting supply and demand constraints lead directly to the question of finding an optimal matching between the ad slots and the advertisers. The problem is inherently online, since we have to show an ad as soon as the request for an ad slot arrives, and we do not have complete information about the arriving ad slots in advance. Furthermore, offline optimization techniques are not even feasible due to the size of the problems, especially given the fact that dealing effectively with the long tail of ad requests is of critical business importance.

The problem of online matching and allocation has generated a lot of interest in the algorithms community, with the introduction of a large number of new problems, models and algorithmic techniques. This is not only due to the importance of the motivation but also due to the new and elegant questions and techniques that emerge. The first objective of this monograph is to provide a systematic survey of this literature.

This theoretical work has had an influential effect on the algorithmic framework used by virtually all of the companies which are in the Internet advertising space. The major contribution has been the introduction of the technique of *bid-scaling*. In this technique we scale the relevant parameter, for example, the bid, by a scaling function, and then choose that edge to match which has the highest scaled bid. This is to be compared to the greedy strategy which simply chooses the edge with the highest bid. The design of optimal algorithms in the online model has also led to the formulation of bid-scaling heuristics. Section 9 provides a brief survey of applications of these algorithms and heuristics, including the domain specific details. Let us quickly note that in the practical problem, there are typically three players in the market: the users of the service, the platform (for example, the search engine), and the advertisers. Thus there are three objective functions to consider: the quality of the ads shown, the revenue to the platform and the return on investment to the advertisers. We will consider these in more detail later, but for most of the survey we will focus on maximizing the efficiency (the total size or weight) of the matching, which can be a good proxy for all relevant objective functions. A second point to note is that different advertising platforms have their own specific settings, for example, second-price auctions vs first-price, single slot vs position auctions, contracts vs auctions, etc. We will abstract these details out for the most part, and mention how they can be modeled, in Section 9.

1.2 Background on Matching: Applications, History and Offline Algorithms

The problem of matching is relevant to a wide variety of important application domains, besides our motivating application of ad allocation. In Economics, matching is relevant whenever there is a two-sided market (see, for example, [86]). One important formulation is the problem of finding a *stable matching* or a *Pareto efficient* matching in a graph [48]. This has found several important applications in the real world: it is used in matching of residents to hospitals (starting with [85]), students to high schools [1], and even kidney donors to recipients (see Kidney Exchanges [84]); Roth and Shapley were awarded the 2012 Nobel Prize in Economics for their influential and impactful work on this topic. Matching, with its generalizations, pervades Computer Science as a core algorithmic problem. For example, in Networking, an important problem is that of finding a good switch scheduling algorithm in input queued (IQ) switch architectures (see [76], among others). This reduces to that of finding a maximum matching to match input ports of a switch to its output ports at every time step. As another example, matching is core to resource allocation problems of various types from the scheduling and Operations Research literature, for example, allocating jobs to machines in cloud computing. Recently, the online matching algorithms from this survey have found applications [54] in crowdsourcing markets.

Besides its high applicability, matching is a central problem in the development of the field of algorithms, and indeed of Theoretical CS. We briefly overview this history next; the rest of this section can be skipped by readers with a strong background in classic matching theory.

The basic algorithms rely on the definition of *augmenting paths*: given a matching M in the graph, an augmenting path is an odd-length (simple) path with its edges alternating between being in M and not, and with the two end edges not in M. Berge's Theorem [17] states that:

Theorem 1.1 (Berge). A matching M is maximum iff it does not admit an augmenting path.

If a matching M admits an augmenting path P, then M can be *augmented* by flipping the membership of the edges of P in and not in M. This transforms M into a matching M' whose size is one more than that of M. An algorithm can proceed in this manner, by starting with any matching, and iteratively finding an augmenting path, and augmenting the matching.

This approach relies on being able to find augmenting paths efficiently. This is possible in bipartite graphs: one can find augmenting paths in bipartite graphs in time $O(|E|)$, by constructing breadth-first search trees (with alternating levels) from unmatched vertices. On bipartite graphs, the problem also has a close relationship with the maximum flow problem; one can reduce unweighted bipartite matching to a max-flow problem by adding a source and a sink to the graph appropriately. The fastest algorithms for this problem [39, 55] run in $O(\sqrt{|V|}|E|)$ time.

The question of finding a maximum matching in general (non-bipartite) graphs is a lot more difficult. Edmonds [42] presented the *Blossom* algorithm to compute a maximum matching in a general graph in polynomial time. The difficulty in general graphs comes precisely due to the presence of odd cycles. The algorithm proceeds by identifying structures, called *blossoms*, with respect to the current matching. A blossom consists of an odd cycle of, say, $2k + 1$ edges, of which exactly k edges belong to the matching, such that there further exists an even length alternating path, called the *stem*, starting with a matched edge at a vertex of the cycle. The algorithm starts with any matching and searches for an augmenting path, which can immediately augment the matching. If it finds a blossom instead, it contracts the blossom into a single vertex and proceeds recursively. If it finds an augmenting path with vertices corresponding to contracted blossoms, then it expands the blossoms (recursively) finding a real augmenting path in the original graph. The running time of this algorithm, with appropriate data structures, is $O(|V|^2|E|)$. The fastest algorithm for matching in general graphs, due to Micali and Vazirani [91], runs in $O(\sqrt{|V|}|E|)$ time.

Let us also quickly note a property of *maximal* matchings, defined as those which cannot be improved upon by only adding more edges.

Theorem 1.2. If M is a maximal matching, and M^* a maximum matching, then $|M| \geq \frac{1}{2}|M^*|$.

This is fairly easy to see: since M is maximal, none of the edges in M^* can be added to it while keeping it a matching. Hence, every edge in M^* uniquely shares an end-point with an edge in M. Thus the number of vertices in M is at least the number of edges in M^*, giving the result. We will generalize this theorem later, to give a bound for greedy online algorithms for all the generalizations of matching that we will study.

In the problem of edge-weighted matching, the edges of G have weights, and the goal is to find a matching with the highest sum of weights of the edges in the matching (in the bipartite case, this is known as the *Assignment Problem*). The algorithm for the edge-weighted bipartite version is more complex than the unweighted problem. It works by updating the matching solution simultaneously with a set of weights on the vertices. This is known as the Hungarian Algorithm [68] (due to Kuhn, based on the work of König and Egerváry), and it is possibly the first example of a primal–dual update algorithm for Linear Programming (here the LP is to maximize the total weight of the matching over the polytope of all fractional matchings, and the weights on the vertices that the algorithm uses correspond to the dual variables of the LP). One observation to make is that all these algorithms are highly offline, that is, not easily adapted to the online setting, a point we will return to in the next section.

As is well-known, there was no fixed formulation of an efficient algorithm at the time that the Blossom algorithm was invented. The Blossom algorithm directly led to the formalization of polynomial time as the correct definition. The impact of this definition is obviously immense to the fields of algorithms, complexity and Computer Science in general, essentially giving us the definition of the complexity class P. Furthermore, the definition of the class $\#P$ is also closely related to matching theory, as Valiant [90] proved that finding the number of perfect matchings in a graph (equivalently, the Permanent of a matrix) is NP-hard, and in fact complete for $\#P$. Matching is also a canonical problem for the study of randomized parallel algorithms and the class RNC; Karp et al. [62], and Mulmuley et al. [82] gave RNC algorithms for finding a maximum matching. The history and algorithms for offline matching have been excellently documented, for example, in the book [71] by Lovász and Plummer.

We will also study the online versions of several generalizations of the basic bipartite matching problem. Most of these are special cases of the Linear Programming problem, which has a vast literature of its own (see, for example, [30]). The classification of these problems and the LP formulations for the offline versions are described in Section 2.

1.3 Online Input

In this monograph we will focus on the online version of the bipartite matching problem and its generalizations. The area of online algorithms and competitive analysis has been very useful in abstracting and studying problems in which the input is not known in advance but is revealed incrementally, even as the algorithm makes its own decisions (see the book by Borodin and El-Yaniv [19]). This is precisely the situation in our motivating applications in which ad slots arrive online, and have to be allocated ads upon arrival, with zero, partial, or stochastic knowledge of the ad slots yet to arrive. We will model our applications via different problems and online input models. In the simplest version of the problem (online bipartite matching), there is a bipartite graph $G(U,V,E)$, in which U is known to the algorithm, vertices in V are unknown, but arrive one at a time, revealing the edges incident on them as they arrive. The algorithm has to match (or forgo) a vertex as soon as it arrives. Furthermore, all matches made are irrevocable;

this is to capture the fact that the arriving vertex v corresponds to an ad-slot on a web page viewed by a user.

Note that all the offline algorithms described in Section 1.2 are "highly offline". They typically involve initialization with some arbitrary matching and subsequent iterative improvements, via augmenting paths or guidance from dual variables. Thus they are not applicable to the online problem where the matches have to be made incrementally as vertices arrive, and are irrevocable. As we will see, the online algorithms work very differently, and often can provide only an approximate solution, that is, with a competitive ratio less than 1.

While our motivation for the online problem comes from ad allocation, large matching questions are becoming more prevalent. Often, the problem is online in nature, for example, the matching of arriving tasks to workers in crowdsourcing applications. Even in applications which are not strictly online, we often face problems with massive data, for example, in a streaming setting. Again, the offline algorithms are not applicable, and we need fast, simple, possibly approximate solutions, for example, in a streaming setting, rather than complex optimal algorithms. We expect that the algorithms surveyed here, or further variants, will be found to be useful in future applications.

Section 2 provides a classification of the different problems and models. Sections 3–8 treats the different problems in detail, giving the different algorithmic techniques. Section 9 describes the application setting and the algorithms and heuristics based on the theoretical results. We will provide open questions throughout the survey, and conclude in Section 10 with a list of additional open problems and future directions.

References

[1] A. Abdulkadiroğlu, P. A. Pathak, and A. E. Roth, "The New York city high school match," *American Economic Review*, pp. 364–367, 2005.

[2] Z. Abrams, O. Mendelevitch, and J. Tomlin, "Optimal delivery of sponsored search advertisements subject to budget constraints," in *ACM Conference on Electronic Commerce*, pp. 272–278, 2007.

[3] M. Adamczyk, "Improved analysis of the greedy algorithm for stochastic matching," *Information Processing Letters*, vol. 111, no. 15, pp. 731–737, 2011.

[4] G. Aggarwal, Y. Cai, A. Mehta, and G. Pierrakos, "Biobjective online bipartite matching," Manuscript.

[5] G. Aggarwal, G. Goel, C. Karande, and A. Mehta, "Online vertex-weighted bipartite matching and single-bid budgeted allocations," in *SODA*, pp. 1253–1264, 2011.

[6] S. Agrawal, Z. Wang, and Y. Ye, "A dynamic near-optimal algorithm for online linear programming," *CoRR,* abs/0911.2974, 2009.

[7] S. Alaei, M. T. Hajiaghayi, V. Liaghat, D. Pei, and B. Saha, "Adcell: Ad allocation in cellular networks," in *ESA*, pp. 311–322, 2011.

[8] N. Andelman and Y. Mansour, "Auctions with budget constraints," in *Scandinavian Workshop on Algorithm Theory (SWAT)*, pp. 26–38, 2004.

[9] J. Aronson, M. Dyer, A. Frieze, and S. Suen, "Randomized greedy matching. II," *Random Structures & Algorithms*, vol. 6, no. 1, pp. 55–73, 1995.

[10] S. Arora, E. Hazan, and S. Kale, "The multiplicative weights update method: A meta-algorithm and applications," *Theory of Computing*, vol. 8, pp. 121–164, 2012.

[11] B. Awerbuch, Y. Azar, and S. Plotkin, "Throughput-competitive on-line routing," in *Proceedings of Annual Symposium on Foundations of Computer Science*, pp. 32–40, 1993.

[12] Y. Azar, B. Birnbaum, A. Karlin, and C. Nguyen, "On revenue maximization in second-price ad auctions," *Algorithms-ESA 2009*, pp. 155–166, 2009.

[13] Y. Azar and A. Litichevskey, "Maximizing throughput in multi-queue switches," *Algorithms–ESA 2004*, pp. 53–64, 2004.

[14] Y. Azar, B. E. Birnbaum, A. R. Karlin, C. Mathieu, and C. T. Nguyen, "Improved approximation algorithms for budgeted allocations," in *ICALP (1)*, pp. 186–197, 2008.

[15] B. Bahmani and M. Kapralov, "Improved bounds for online stochastic matching," in *ESA (1)*, pp. 170–181, 2010.

[16] N. Bansal, A. Gupta, J. Li, J. Mestre, V. Nagarajan, and A. Rudra, "When lp is the cure for your matching woes: Improved bounds for stochastic matchings — (extended abstract)," in *ESA (2)*, pp. 218–229, 2010.

[17] C. Berge, "Two theorems in graph theory," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 43, no. 9, p. 842, 1957.

[18] B. E. Birnbaum and C. Mathieu, "On-line bipartite matching made simple," *SIGACT News*, vol. 39, no. 1, pp. 80–87, 2008.

[19] A. Borodin and R. El-Yaniv, *Online Computation and Competitive Analysis*, volume 53. Cambridge: Cambridge University Press, 1998.

[20] N. Buchbinder, M. Feldman, A. Ghosh, and J. Naor, "Frequency capping in online advertising," in *WADS*, pp. 147–158, 2011.

[21] N. Buchbinder, K. Jain, and J. Naor, "Online primal-dual algorithms for maximizing ad-auctions revenue," in *ESA*, pp. 253–264, 2007.

[22] N. Buchbinder and J. Naor, "The design of competitive online algorithms via a primal-dual approach," *Foundations and Trends in Theoretical Computer Science*, vol. 3, no. 2–3, pp. 93–263, 2009.

[23] D. Chakrabarty and G. Goel, "On the approximability of budgeted allocations and improved lower bounds for submodular welfare maximization and gap," in *FOCS*, pp. 687–696, 2008.

[24] T. Chakraborty, E. Even-Dar, S. Guha, Y. Mansour, and S. Muthukrishnan, "Selective call out and real time bidding," *Internet and Network Economics*, pp. 145–157, 2010.

[25] D. Charles, D. Chakrabarty, M. Chickering, N. R. Devanur, and L. Wang, "Budget smoothing for internet ad auctions: a game theoretic approach," in *Proceedings of the ACM Conference on Electronic Commerce*, pp. 163–180, 2013.

[26] D. X. Charles, M. Chickering, N. R. Devanur, K. Jain, and M. Sanghi, "Fast algorithms for finding matchings in lopsided bipartite graphs with applications to display ads," in *ACM Conference on Electronic Commerce*, pp. 121–128, 2010.

[27] C. Chekuri and S. Khanna, "A polynomial time approximation scheme for the multiple knapsack problem," *SIAM Journal on Computing*, vol. 35, no. 3, pp. 713–728, 2005.

[28] N. Chen, N. Immorlica, A. R. Karlin, M. Mahdian, and A. Rudra, "Approximating matches made in heaven," in *ICALP (1)*, pp. 266–278, 2009.
[29] Y. Chen, P. Berkhin, B. Anderson, and N. R. Devanur, "Real-time bidding algorithms for performance-based display ad allocation," in *KDD*, pp. 1307–1315, 2011.
[30] V. Chvátal, *Linear Programming*. WH Freeman, 1983.
[31] K. P. Costello, P. Tetali, and P. Tripathi, "Stochastic matching with commitment," in *Automata, Languages, and Programming*, pp. 822–833, Springer, 2012.
[32] N. R. Devanur and K. Jain, "Online matching with concave returns," in *Proceedings of the Symposium on Theory of Computing*, pp. 137–144, 2012.
[33] N. Devanur, K. Jain, and R. Kleinberg, "Randomized primal-dual analysis of ranking for online bipartite matching," in *SODA*, 2013.
[34] N. R. Devanur, "Online algorithms with stochastic input," *SIGecom Exchanges*, vol. 10, no. 2, pp. 40–49, 2011.
[35] N. R. Devanur and T. P. Hayes, "The adwords problem: online keyword matching with budgeted bidders under random permutations," in *ACM Conference on Electronic Commerce*, pp. 71–78, 2009.
[36] N. R. Devanur, K. Jain, B. Sivan, and C. A. Wilkens, "Near optimal online algorithms and fast approximation algorithms for resource allocation problems," in *ACM Conference on Electronic Commerce*, pp. 29–38, 2011.
[37] N. R. Devanur, B. Sivan, and Y. Azar, "Asymptotically optimal algorithm for stochastic adwords," in *Proceedings of the ACM Conference on Electronic Commerce*, pp. 388–404, 2012.
[38] N. B. Dimitrov and C. G. Plaxton, "Competitive weighted matching in transversal matroids," in *ICALP (1)*, pp. 397–408, 2008.
[39] E. A. Dinic, "Algorithm for solution of a problem of maximum flow in networks with power estimation," *In Soviet Math. Dokl*, vol. 11, no. 5, no. 5, pp. 1277–1280, 1970.
[40] M. E. Dyer and A. M. Frieze, "Randomized greedy matching," *Random Struct. Algorithms*, vol. 2, no. 1, pp. 29–46, 1991.
[41] E. B. Dynkin, "The optimum choice of the instant for stopping a markov process," *Soviet Mathematics. Doklady*, vol. 4, 1963.
[42] J. Edmonds, "Paths, trees, and flowers," *Canadian Journal of mathematics*, vol. 17, no. 3, pp. 449–467, 1965.
[43] U. Feige and J. Vondrak, "Approximation algorithms for allocation problems: Improving the factor of 1-1/e," in *Annual IEEE Symposium on Foundations of Computer Science, 2006. FOCS'06*, pp. 667–676, 2006.
[44] J. Feldman, M. Henzinger, N. Korula, V. S. Mirrokni, and C. Stein, "Online stochastic packing applied to display ad allocation," in *ESA (1)*, pp. 182–194, 2010.
[45] J. Feldman, N. Korula, V. S. Mirrokni, S. Muthukrishnan, and M. Pál, "Online ad assignment with free disposal," in *WINE*, pp. 374–385, 2009.
[46] J. Feldman, A. Mehta, V. S. Mirrokni, and S. Muthukrishnan, "Online stochastic matching: Beating 1-1/e," in *FOCS*, pp. 117–126, 2009.

[47] L. Fleischer, M. X. Goemans, V. S. Mirrokni, and M. Sviridenko, "Tight approximation algorithms for maximum general assignment problems," in *Proceedings of the Annual ACM-SIAM symposium on Discrete Algorithm*, pp. 611–620, 2006.

[48] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *American Mathematical Monthly*, pp. 9–15, 1962.

[49] R. Garg, V. Kumar, and V. Pandit, "Approximation algorithms for budget-constrained auctions," *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques*, pp. 102–113, 2001.

[50] G. Goel and A. Mehta, "Adwords auctions with decreasing valuation bids," in *WINE*, pp. 335–340, 2007.

[51] G. Goel and A. Mehta, "Online budgeted matching in random input models with applications to adwords," in *SODA*, pp. 982–991, 2008.

[52] M. Goemans and J. Kleinberg, "An improved approximation ratio for the minimum latency problem," *Mathematical Programming*, vol. 82, no. 1, pp. 111–124, 1998.

[53] B. Haeupler, V. Mirrokni, and M. Zadimoghaddam, "Online stochastic weighted matching: Improved approximation algorithms," *Internet and Network Economics*, pp. 170–181, 2011.

[54] C.-J. Ho and J. W. Vaughan, "Online task assignment in crowdsourcing markets," in *AAAI*, 2012.

[55] J. E. Hopcroft and R. M. Karp, "An nˆ5/2 algorithm for maximum matchings in bipartite graphs," *SIAM Journal on Computing*, vol. 2, no. 4, pp. 225–231, 1973.

[56] P. Jaillet and X. Lu, "Online stochastic matching: New algorithms with better bounds," Manuscript.

[57] K. Jain, M. Mahdian, E. Markakis, A. Saberi, and V. V. Vazirani, "Greedy facility location algorithms analyzed using dual fitting with factor-revealing lp," *Journal of ACM*, vol. 50, no. 6, pp. 795–824, 2003.

[58] B. Kalyanasundaram and K. Pruhs, "An optimal deterministic algorithm for online b-matching," *Theoretical Computer Science*, vol. 233, no. 1–2, pp. 319–325, 2000.

[59] M. Kapralov, I. Post, and J. Vondrák, "Online and stochastic variants of welfare maximization," *CoRR,* abs/1204.1025, 2012.

[60] C. Karande, A. Mehta, and R. Srikant, "Optimization of budget constrained spend in search advertising," in *WSDM*, 2013.

[61] C. Karande, A. Mehta, and P. Tripathi, "Online bipartite matching with unknown distributions," in *STOC*, pp. 587–596, 2011.

[62] R. M. Karp, E. Upfal, and A. Wigderson, "Constructing a perfect matching is in random nc," *Combinatorica*, vol. 6, no. 1, pp. 35–48, 1986.

[63] R. M. Karp, U. V. Vazirani, and V. V. Vazirani, "An optimal algorithm for on-line bipartite matching," in *STOC*, pp. 352–358, 1990.

[64] T. Kesselheim, K. Radke, A. Tönnis, and B. Vöcking, "An optimal online algorithm for weighted bipartite matching and extensions to combinatorial auctions," in *Algorithms–ESA 2013*, pp. 589–600, Springer, 2013.

[65] S. Khot, R. J. Lipton, E. Markakis, and A. Mehta, "Inapproximability results for combinatorial auctions with submodular utility functions," *Algorithmica*, vol. 52, no. 1, pp. 3–18, 2008.
[66] N. Korula, V. Mirrokni, and M. Zadimoghaddam, "Bicriteria online matching: Maximizing weight and cardinality," Manuscript.
[67] N. Korula and M. Pál, "Algorithms for secretary problems on graphs and hypergraphs," in *ICALP (2)*, pp. 508–520, 2009.
[68] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 2006.
[69] B. Lehmann, D. J. Lehmann, and N. Nisan, "Combinatorial auctions with decreasing marginal utilities," *Games and Economic Behavior*, vol. 55, no. 2, pp. 270–296, 2006.
[70] D. V. Lindley, "Dynamic programming and decision theory," *Applied Statistics*, pp. 39–51, 1961.
[71] L. Lovász and M. D. Plummer, *Matching Theory,* volume 367. American Mathematical Society, 2009.
[72] M. Mahdian, H. Nazerzadeh, and A. Saberi, "Allocating online advertisement space with unreliable estimates," in *ACM Conference on Electronic Commerce*, pp. 288–294, 2007.
[73] M. Mahdian and Q. Yan, "Online bipartite matching with random arrivals: an approach based on strongly factor-revealing lps," in *STOC*, pp. 597–606, 2011.
[74] V. H. Manshadi, S. O. Gharan, and A. Saberi, "Online stochastic matching: Online actions based on offline statistics," in *SODA*, pp. 1285–1294, 2011.
[75] R. McEliece, E. Rodemich, H. Rumsey, and L. Welch, "New upper bounds on the rate of a code via the delsarte-macwilliams inequalities," *IEEE Transactions on Information Theory*, vol. 23, no. 2, pp. 157–166, 1977.
[76] N. McKeown, A. Mekkittikul, V. Anantharam, and J. Walrand, "Achieving 100% throughput in an input-queued switch," *IEEE Transactions on Communications*, vol. 47, no. 8, pp. 1260–1267, 1999.
[77] A. Mehta and D. Panigrahi, "Online matching with stochastic rewards," in *IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, 2012.
[78] A. Mehta, A. Saberi, U. V. Vazirani, and V. V. Vazirani, "Adwords and generalized online matching," *Journal of ACM*, vol. 54, no. 5, 2007.
[79] V. S. Mirrokni, S. O. Gharan, and M. Zadimoghaddam, "Simultaneous approximations for adversarial and stochastic online budgeted allocation," in *SODA*, pp. 1690–1701, 2012.
[80] V. S. Mirrokni, M. Schapira, and J. Vondrák, "Tight information-theoretic lower bounds for welfare maximization in combinatorial auctions," in *ACM Conference on Electronic Commerce*, pp. 70–77, 2008.
[81] R. Motwani, R. Panigrahy, and Y. Xu, "Fractional matching via balls-and-bins," in *APPROX-RANDOM*, pp. 487–498, 2006.
[82] K. Mulmuley, U. V. Vazirani, and V. V. Vazirani, "Matching is as easy as matrix inversion," in *Proceedings of the Annual ACM Symposium on Theory of Computing*, pp. 345–354, 1987.
[83] S. Muthukrishnan, "Ad exchanges: Research issues," *Internet and Network Economics*, pp. 1–12, 2009.

[84] A. Roth, T. Sonmez, and U. Unver, "Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences," *American Economic Review*, vol. 97, no. 3, 2007.

[85] A. E. Roth, "The evolution of the labor market for medical interns and residents: a case study in game theory," *The Journal of Political Economy*, pp. 991–1016, 1984.

[86] A. E. Roth and M. A. O. Sotomayor, *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis,* vol. 18. Cambridge University Press, 1992.

[87] A. E. Roth, T. Sonmez, and M. U. Unver, "Pairwise kidney exchange," *Journal of Economic Theory*, vol. 125, pp. 151–188, 2005.

[88] D. B. Shmoys and É. Tardos, "An approximation algorithm for the generalized assignment problem," *Math. Program*, vol. 62, pp. 461–474, 1993.

[89] A. Srinivasan, "Budgeted allocations in the full-information setting," in *APPROX-RANDOM*, pp. 247–253, 2008.

[90] L. G. Valiant, "The complexity of computing the permanent," *Theoretical Computer Science*, vol. 8, no. 2, pp. 189–201, 1979.

[91] V. V. Vazirani, "A theory of alternating paths and blossoms for proving correctness of the general graph maximum matching algorithm," *Combinatorica*, vol. 14, no. 1, pp. 71–109, 1994.

[92] E. Vee, S. Vassilvitskii, and J. Shanmugasundaram, "Optimal online assignment with forecasts," in *ACM Conference on Electronic Commerce*, pp. 109–118, 2010.

[93] J. Vondrák, "Optimal approximation for the submodular welfare problem in the value oracle model," in *STOC*, pp. 67–74, 2008.