Theoretical Computer Science 621 (2016) 82–91

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Maximum bipartite matchings with low rank data: Locality and perturbation analysis

Xingwu Liu [a,1], Shang-Hua Teng [b,*,2]

[a] Institute of Computing Technology, Chinese Academy of Sciences, China
[b] Department of Computer Science, Viterbi School of Engineering, USC, United States

A R T I C L E   I N F O

Article history:
Received 14 February 2015
Accepted 26 January 2016
Available online 2 February 2016
Communicated by J. Díaz

Keywords:
Bipartite graphs
Maximum matching
Matrix decomposition
Permutation

A B S T R A C T

The maximum-weighted bipartite matching problem between two sets $U = V = [1:n]$ is defined by a matrix $\mathbf{W} = (w_{ij})_{n\times n}$ of "affinity" data. Its goal is to find a permutation π over $[1:n]$ whose total weight $\sum_i w_{i,\pi(i)}$ is maximized. In various practical applications,[3] the affinity data may be of low rank (or approximately low rank): we say $\mathbf{W}$ has *rank* at most r if there are $2r$ vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{v}_1, \ldots \mathbf{v}_r \in \mathbb{R}^n$ such that

$$W = \sum_{i=1}^{r} \mathbf{u}_i \mathbf{v}_i^T.$$

In this paper, we partially address a question raised by David Karger who asked for a characterization of the structure of the maximum-weighted bipartite matchings when the rank of the affinity data is low. In particular, we study the following locality property: For an integer $k > 0$, we say that the bipartite matchings of G have *locality* at most k if for every sub-optimal matching π of G, there exists a matching σ of larger total weight that can be reached from π by an augmenting cycle of length at most k.

We prove the following main theorem: For every $W \in [0,1]^{n\times n}$ of rank r and $\epsilon \in [0,1]$, there exists $\tilde{W} \in [0,1]^{n\times n}$ such that (i) $\tilde{W}$ has rank at most $r+1$, (ii) the entry-wise ∞-norm $\|W - \tilde{W}\|_\infty \le \epsilon$, and (iii) the weighted bipartite graph with affinity data $\tilde{W}$ has locality at most $\lceil r/\epsilon \rceil^r$. In contrast, this property is not true if perturbations are not allowed. We also give a tight bound for the binary case.

1. Introduction

The maximum-weighted bipartite matching [4,7] is a fundamental problem at the intersection of graph theory, geometric design, and combinatorial optimization [10,5]. An instance of this optimization problem is a weighted bipartite graph $G = (U, V, W)$, where $U = V = [1:n] \triangleq \{1, 2, \ldots n\}$ denote the two vertex sets, and the matrix $W = (w_{i,j})_{n\times n} \in \mathbb{R}^{n\times n}$ defines

* Corresponding author.
E-mail addresses: Xingwuliu@gmail.com (X. Liu), shanghua@usc.edu (S.-H. Teng).

1 The work was done when the author was visiting USC. The work was partially supported by National Natural Science Foundation of China (61173009).
2 Supported in part by the NSF grants 1111270 and 0964481.

3 An example will be given in Section 1.

http://dx.doi.org/10.1016/j.tcs.2016.01.033
0304-3975/

the affinity weights[4] between every pair of elements between U and V. The optimization objective is to find a perfect matching, which can be expressed by a permutation π over $[1:n]$, such that the total weight of the matching

$$\mu_W(\pi) = \sum_{i=1}^{n} w_{i,\pi(i)}$$

is maximized.

In various practical applications, the affinity data may be of low rank (or approximately low rank). As an example, we consider the following basic problem of assigning university volunteers to mentor high-school students in an outreach program: Assume n university students volunteer to teach a summer program at one of the n high schools in the city of Los Angeles. The city would like to assign each high school exactly one mentor. However, the teaching capacities $\mathbf{c} = (c_1, \ldots, c_n)$ vary from mentor to mentor and the educational needs $\mathbf{d} = (d_1, \ldots, d_n)$ vary from school to school. Suppose the educational benefit of a particular assignment π can be (approximately) expressed by $\sum_{i=1}^{n} c_i \cdot d_{\pi(i)}$. To maximize this potential benefit of mentorships, it amounts to find a maximum weighted matching in a bipartite graph where $W = \mathbf{c} \cdot \mathbf{d}^T$ is a matrix of rank 1. In general, the needs of the school could be multidimensional, and are characterized by a small number of principal concerns. For instance, the educational needs maybe approximately classified by the "academic needs" and "social needs". The mentoring capacities of the volunteers may also be characterized according to these two dimensions. In this context, the educational needs and mentoring capacities are each approximated by two vectors

$$\mathbf{d}^a = (d_1^a, \ldots, d_n^a) \quad \mathbf{d}^s = (d_1^s, \ldots, d_n^s)$$
$$\mathbf{c}^a = (c_1^a, \ldots, c_n^a) \quad \mathbf{c}^s = (c_1^s, \ldots, c_n^s),$$

and the educational benefit of a particular assignment π is approximately expressed by

$$\sum_{i=1}^{n} (c_i^a \cdot d_{\pi(i)}^a + c_i^s \cdot d_{\pi(i)}^s).$$

To maximize this benefit, it amounts to find a maximum weighted matching in a bipartite graph where $W = \mathbf{c}^a \cdot (\mathbf{d}^a)^T + \mathbf{c}^s \cdot (\mathbf{d}^s)^T$ is an affinity matrix of rank 2.

Partially motivated by this type of applications, David Karger [6] (in 2009) asked to characterize the structure of the bipartite matchings when the affinity data is of low rank. For an integer r, W has rank at most r if there exist $2r$ vectors $\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{v}_1, \ldots \mathbf{v}_r \in \mathbb{R}^n$ such that

$$W = \sum_{i=1}^{r} \mathbf{u}_i \mathbf{v}_i^T.$$

A combinatorial study of algorithmic interest is about the structure of the augmenting cycles for improving sub-optimal bipartite matchings. This subject will be the focus of our work as well. Below, we will use $\overrightarrow{i_0 i_1 \ldots i_{k-1}}$ to denote the cycle (i.e., the cyclic permutation) that maps i_l to $i_{l+1 \bmod k}$, for $0 \le l \le k-1$. The number k here is called the *length* of the cycle. It is well known that when a matching π is not maximum with respect to W, there exists a cycle $\sigma = \overrightarrow{i_0 i_1 \ldots i_{k-1}} \in S_n$ such that the composite permutation $\sigma \circ \pi \in S_n$ improves π. Such an σ is referred to as an *augmenting cycle of* π (with respective to W).

Before stating the initial conjecture that motivated our work, we first consider the case when $r = 1$, i.e., there are real vectors $\mathbf{u} = (u_1, \ldots, u_n)^T$ and $\mathbf{v} = (v_1, \ldots, v_n)^T$ such that $W = \mathbf{u}\mathbf{v}^T$. We say that a permutation σ *sorts the entries of* $\mathbf{u}$ *in descending order* if

$$u_{\sigma(1)} \ge u_{\sigma(2)} \ge \cdots \ge u_{\sigma(n)}.$$

Then, it is relatively standard (e.g., see [2]) to show that π is a maximum-weighted bipartite matching if and only if the following is true: Suppose σ is the permutation that sorts the entries of $\mathbf{u}$ in descending order, then $\pi \circ \sigma$ also sorts the entries of $\mathbf{v}$ in descending order. Thus, if π is sub-optimal with respect to a rank-one affinity matrix W, then π has an augmenting cycle of length 2.

In general, let us define the following:

Definition 1.1 *(Locality of weighted bipartite matchings).* The locality of an affinity matrix $W \in \mathbb{R}^{n \times n}$ is defined to be the minimum k such that every sub-optimal matching with respect to W has an augmenting cycle of length at most k.

[4] One usually assumes that the affinity weights are nonnegative, which implies that every maximum weighted matching can be extended to a perfect matching between U and V with the same total weights. However, by explicitly requiring perfect matchings, one can apply translation to any affinity matrix with negative entries to define an equivalent nonnegative instance.