THE PROBABILISTIC ANALYSIS OF A HEURISTIC FOR THE ASSIGNMENT PROBLEM

David Avis

*C. W. Lai**

School of Computer Science

McGill University

805 Sherbrooke St. W.

Montreal, Canada, H3A 2K6

ABSTRACT

We present a heuristic to solve the m by m assignment problem in $O(m^2)$ time. The assignment problem is formulated as a weighted complete bipartite graph $G = (S, T, E)$, $|S| = |T| = m$. For convenience we assume that m is even. The main procedure in the heuristic is to construct a graph $G_d = (S, T, E_d)$ which is a subgraph of G , $|E_d| = 4dn$, $n = \frac{m}{2}$, such that we can find a perfect matching in G_d with probability at least $1 - \frac{1}{3}\left(\frac{d}{n}\right)^{d^2-4d-1}$. An $O(|S||E|)$ exact algorithm is used to find a minimum weight matching M in G_d. Any unmatched vertices in G relative to M are then matched by a greedy algorithm. The expected value of the total cost of the matching found by the heuristic is shown to be less than 6 if the costs are independent and identically distributed uniformly in the unit interval. Further, with the above probability, the heuristic produces a solution which is at most 6 times the optimal solution.

Key Words. Assignment Problem, Heuristics, Probabilistic Analysis of Algorithms, Weighted Matchings.

1. Introduction

Let $G = (S, T, E)$ be the complete bipartite graph with vertex sets S and T each of cardinality m, and E the set of edges between S and T. Associated with each edge (i, j) there is a weight $c_{i,j}$. A matching M in G is a set of edges such that no two edges in M are incident to the same vertex of G. The matching is perfect if each vertex in G is adjacent to some edge in M. A minimum weight perfect

* Current Address: G.P.O. 2706, Hong Kong

matching for G is a perfect matching for which the sum of the edge weights is minimum. The assignment problem is to find such a minimum weight matching.

There are exact algorithms to solve the assignment problem in $O(m^3)$ time[10]. Karp[7] has found an algorithm that finds an optimal solution in expected time $O(n^2 \log n)$. Several heuristics exist for the assignment problem [8], [6], [1] [3]. A survey of heuristics for the weighted matching problem and some applications is contained in[2]. In order to evaluate the solutions produced by these heuristics, it is natural to ask what is the expected size of the optimal solution, given a probability distribution for the edge weights. If the edge weights are independently distributed on the unit interval, Walkup[12] has proved that this expected value is bounded above by the constant 3, for all m, a very remarkable result. The idea behind Walkup's proof is the construction of a sparse subgraph of G in which the edges all have low weight. He defines a subgraph G_d, for some small integer d, by choosing d edges randomly from each vertex in S and T. In[13] it is shown that with probability approaching one as m tends to infinity, this graph has a perfect matching. The natural candidate for such a subgraph would be to choose for each vertex, the d smallest weight edges that are incident with it. Unfortunately this will not work, since the edges chosen in such a way will not be independent. Consider, for example, the minimum weight edge in G. It will be chosen twice, once by a vertex in S and once by a vertex in T. Walkup avoids this problem by using a probabilistic trick: each edge is replaced by two directed edges with opposite directions. The weights on the directed edges are chosen from distributions such that the minimum of the two edge weights still has uniform distribution on the unit interval. This is sufficient for proving the result stated, but does not provide a constructive method that can be used to produce a heuristic.

In this paper we avoid the dependency problem by constructing a different kind of random graph. We prove analogues of Walkup's results for this new graph. It is shown to almost surely have a perfect matching, and it is shown how to construct this graph directly from the given edge weights. We prove that the perfect matching thus found has expected weight at most 6. The heuristic runs in $O(m^2)$ time and produces a solution which with probability approaching one, is at most 6 times the optimum value. This is the only known heuristic with a constant ratio bound, with probability approaching one. In section 2 the heuristic is presented along with a complexity analysis. In section 3 we show that the random graph constructed by the heuristic has a perfect matching with probability approaching one. This is followed in section 4 by a proof that the expected weight of the matching produced is less than 6. The results given in this paper were first published in the second author's Master's thesis[9]. This thesis also contains extensive numerical results.

2. The Heuristic

For convenience we assume that m is even and set $n = m/2$. The heuristic is supplied with an integer parameter d which must be at least 5. This parameter controls the quality of the solution and the cost of the computation. In practice, setting $d = 5$ appears to be most satisfactory. The algorithm consists of three steps:

- The construction of a sparse directed random graph G_d with $4dn$ edges;
- The solution of the maximum cardinality minimum weight matching problem in the undirected graph obtained by ignoring the directions on edges of G_d;
- The matching of any unmatched vertices in G_d using a greedy heuristic.

Step-1.(Construction of a directed graph $G_d = (S_1 \cup S_2, T_1 \cup T_2, E_d)$ with $4dn$ edges, see Figure 1.)

Partition S and T into subsets S_1, S_2 and T_1, T_2 each of cardinality n.

For $v \in S_i, i = 1, 2$

$$N(v) = \{\ w \in T_i | c_{v,w} \textit{ is one of the } d \textit{ smallest edge weights incident } v\ \}.$$

For $w \in T_i, i = 1, 2$

$$N(w) = \{\ v \in S_{3-i} | c_{v,w} \textit{ is one of the } d \textit{ smallest edge weights incident } w\ \}.$$

Set $E_d = \{(v, w) \mid v \in S_i \cup T_i,\ w \in N(v),\ i = 1, 2\ \}$.

Step-2.(Find a minimum weight matching in G_d.)

Apply an $O(|S||E|)$ minimum weight matching algorithm to G_d ignoring the directions on the edges. Let M_d be the resulting matching.

Step-3.(Match the unmatched vertices in G_d by a greedy algorithm.)

$$S^* = \{\ \textit{unmatched vertices in } S \textit{ relative } M_d\ \}.$$

$$T^* = \{\ \textit{unmatched vertices in } T \textit{ relative } M_d\ \}.$$

while $|\, S^* \,| \geq 1$ **do**

Choose a vertex $s \in S^*$ and find a vertex $t \in T^*$ such that

$$c_{s,t} = \min\{c_{s,v} \mid v \in T^*\}.\ \ S^* = S^* - s;\ \ T^* = T^* - t;\ \ M_d = M_d \bigcup (s,t).$$

end.

Step 1 involves at most $O(2d|S|)$ operations; step 2 needs $O(|S||E_d|) = O(|S|^2)$ operations; and step 3 requires at most $O(|S|^2)$ operations. The memory space required by G and G_d is $O(|S|^2)$ and $O(2d|S|)$ respectively. There are exact algorithms for step 2 that require $O(|S|^2)$ space. Therefore the overall time and space requirements for the algorithm are $O\ (m^2)$.

3. Probability of a Perfect Matching in G_d .

In this section we show that with probability approaching one, the graph $G_d = (S, T, E_d)$ constructed by the heuristic contains a perfect matching. It is convenient to consider the following random graph model for a random $2n\ x\ 2n$ directed bipartite graph with vertex sets $(S_1, S_2; T_1, T_2)$, $|S_i| = |T_i| = n,\ i = 1, 2.$ For $i = 1, 2$ and each vertex $v \in S_i$, select d neighbours at random in T_i. For $i = 1, 2$ and each vertex $W \in T_i$, select d neighbours at random in S_{3-i}. A graph constructed in such a way will arise with equal probability by an application of Step 1 of the heuristic. This follows from the fact that selecting the d smallest from a set of n independent and identically distributed random variables gives a uniform random subset of d of the random variables. In this section, we deal with the equivalent formulation mentioned above which does not involve edge weights. Finally, in Section 4, we reintroduce the edge weights to prove our main result. We will apply the Konig-Hall theorem, see for example Bondy and Murty[4], to show that G_d almost surely has a perfect matching. We require the following definitions.

Let k be a positive integer, satisfying $d + 1 \leq k \leq 2n - d$, let A be a subset of T with cardinality $k - 1$ and let B be a subset of S with cardinality k . Set

$$A_i = T_i \cap A \qquad a_i = |A_i| \quad i = 1, 2$$

$$B_i = S_i \cap B \qquad b_i = |B_i| \quad i = 1, 2.$$

For any subset U of vertices, let

$$\Gamma(U) = \{w : (v, w) \in E_d \text{ for } some\ v \in U\}.$$

We say that (A, B) is a *blocking* $k - pair$ in G_d if the following hold (see Figure 2):

$$\Gamma(B_i) \subseteq A_i, \qquad \Gamma(T_i - A_i) \subseteq S_{3-i} - B_{3-i}, \quad i = 1, 2.$$

It can be seen from the construction that certain values of a_i place restrictions on the corresponding values of b_i. In particular:

(i) If $a_i < d$ then $b_i = 0$;

(ii) If $a_i = 0$ then $d + 1 \leq k \leq n - d$;

(iii) If $a_1, a_2 \geq d$ then $b_1, b_2 \leq n - d$; and

(iv) $a_1 + a_2 = k - 1, \quad b_1 + b_2 = k$.

We say that the 4-tuple is *valid* if it is consistent with (i)-(iv) above.

By the Konig-Hall theorem, G_d has a perfect matching if and only if it had no blocking k –pair. Let $P^{(d)}{}_{a_1,b_1}(k)$ be the probability that (A, B) is a blocking k-pair in G_d . This notation is justified by observing that the probability of the algorithm creating a blocking k-pair (A, B) is determined only by the cardinalities a_1, a_2, b_1, b_2. Since a_2 and b_2 are determined once a_1 and b_1 are fixed, we omit them as subscripts of P. Then if a_1, b_1 satisfy the above conditions for a blocking k-pair,

$$P^{(d)}{}_{a_1,b_1}(k) = \left[\frac{\binom{a_1}{d}}{\binom{n}{d}}\right]^{b_1} \left[\frac{\binom{a_2}{d}}{\binom{n}{d}}\right]^{b_2} \left[\frac{\binom{n-b_2}{d}}{\binom{n}{d}}\right]^{n-a_1} \left[\frac{\binom{n-b_1}{d}}{\binom{n}{d}}\right]^{n-a_2} . \tag{1}$$

We denote by $\beta^{(d)}(k)$ the expected number of blocking k –pairs. Then we have

$$\beta^{(d)}(k) = \sum_{b_1+b_2=k} \sum_{a_1+a_2=k-1} \binom{n}{a_1}\binom{n}{a_2}\binom{n}{b_1}\binom{n}{b_2} \; P^{(d)}{}_{a_1,b_1}(k) . \tag{2}$$

Our main result is the following. Let $\beta^{(d)}$ denote the expected number of blocking pairs in G_d.

Theorem 1.

$$For\ n \geq 2d \geq 10 \qquad \beta^{(d)} < \frac{1}{3}\left(\frac{d}{n}\right)^{d^2-4d-1}.$$

The theorem proves the claim that G_d almost certainly has a perfect matching, since by the Markov inequality,

$$P\{G_d has\ no\ perfect\ matching\ \} \leq \beta^{(d)}.$$

In order to prove this result we need a series of combinatorial lemmas.

Lemma 1 (See[11].)

$For\ positive\ integers\ k,\ m\ with\ \ k < m,$

$$(a) \qquad \binom{m}{k} \leq \frac{1}{\sqrt{2\pi k}}\left(\frac{em}{k}\right)^k,$$

$$(b) \qquad \binom{m}{k} \leq \frac{m^m}{k^k\ \ (m-k)^{(m-k)}}. \quad \square$$

Lemma 2.

$$\frac{\binom{p}{j}}{\binom{q}{j}} \leq \left(\frac{p}{q}\right)^j, \qquad 0 \leq j \leq p \leq q. \quad \square$$

Lemma 3. For $2 \leq d+1 \leq k < n$, every valid 4-tuple (a_1, a_2, b_1, b_2) satisfies:

$$(a) \quad \left(\frac{a_1}{n}\right)^{b_1} \left(\frac{a_2}{n}\right)^{b_2} \leq \left(\frac{k-1}{n}\right)^k;$$

$$(b) \quad \left(\frac{n-b_2}{n}\right)^{n-a_1} \left(\frac{n-b_1}{n}\right)^{n-a_2} \leq \left(\frac{n-k}{n}\right)^{n-k}$$

$$(c) \quad \left(\frac{a_1}{n}\right)^{b_1}\left(\frac{a_2}{n}\right)^{b_2}\left(\frac{n-b_2}{n}\right)^{n-a_1}\left(\frac{n-b_1}{n}\right)^{n-a_2} \leq \left(\frac{d}{n}\right)^d.$$

(*d*) *If in addition* $k \le n - d$ *then*

$$\left(\frac{k-1}{n}\right)^k \left(\frac{n-k}{n}\right)^{n-k} \le \left(\frac{d}{n}\right)^d.$$

Proof. The various parts of the lemma are proved in more or less the same way. We just outline their proof, leaving the details to the interested reader.

(a) This is an immediate consequence of the arithmetic-geometric mean inequality, see[11], p. 76.

(b) Set $a = a_2$, $b = b_1$, $a_1 = k - a - 1$, and $b_2 = k - b$. We first assume that $a_1, a_2 \ge d$. Then, using elementary calculus,

$$\begin{aligned} n^{-1}(n-b_1)^{n-a_2}(n-b_2)^{n-a_1} &\le (n-b)^{n-a}(n-k+b)^{n-k+a} \\ &\le (n-a)^{n-a}(n-k+a)^{n-k+a} \\ &\le (n-d)^{n-d}(n-k+d)^{n-k+d} \\ &\le n^n(n-k)^{n-k}. \end{aligned}$$

Now, suppose that $a_1 < d$. Since a_1, a_2, b_1, b_2 is valid, $b_1 = 0$, hence $b_2 = k$. Then the left hand side of b reduces to

$$\left(\frac{n-k}{n}\right)^{n-a_1} \le \left(\frac{n-k}{n}\right)^{n-d+1} \le \left(\frac{n-k}{n}\right)^{n-k}$$

as required. The case $a_2 < d$ is symmetric.

(c) Let h denote the left hand side. Then from the third inequality in the proof of (b), part (a) and elementary calculus, we obtain

$$h \le \left(\frac{k}{n}\right)^k \left(\frac{n-k+d}{n}\right)^{n-k+d} \le \left(\frac{d}{n}\right)^d.$$

(d) Follows from elementary calculus. □

The following rather technical lemma will be required in the proof of Lemma 6, but can be skipped without loss of continuity.

Lemma 4.

$$\textit{If}\quad 1 \le \frac{n}{4} \le \frac{k}{2} \le a < k-1 \ \text{ and }\ a < b \le k \le n \quad \textit{then}$$

$$h = a^{b-a-1}\,(k-1-a)^{a-b+1}\,(n-b)^{\,(a+b)-k}\,(n-k+b)^{k-(a+b)} \ \le\ 1.$$

Proof. Using standard inequalities for logarithms[11] one obtains

$$\frac{1}{2}\log\left(\frac{a}{k-1-a}\right) \le \frac{2a-k+1}{k-1} + \frac{1}{3}\,\frac{\left(\dfrac{2a-k+1}{k-1}\right)^3}{1-\left(\dfrac{2a-k+1}{k-1}\right)^2}\,,$$

and

$$\frac{1}{2}\log\left(\frac{n-k+b}{n-b}\right) \ge \frac{2b-k}{2n-k}\ .$$

A elementary calculation shows that

$$(b-a-1)\,\frac{2a-k+1}{k-1} - (a+b-k)\,\frac{2b-k}{2n-k} \le -\,\frac{(2a-k+1)(2a-k+2)}{3k}\ .$$

Therefore

$$\frac{1}{2}\log h \le \frac{1}{2}\left[(b-a-1)\log(\frac{a}{k-1-a}) - (a+b-k)\log(\frac{n-k+b}{n-b})\right]$$

$$\le (b-a-1)\left[\frac{2a-k+1}{k-1} + \frac{1}{3}\,\frac{\left(\dfrac{2a-k+1}{k-1}\right)^3}{1-\left(\dfrac{2a-k+1}{k-1}\right)^2}\right] - (a+b-k)\,\frac{2b-k}{2n-k}$$

$$\le\ -\,\frac{(2a-k+1)(2a-k+2)}{3k} + \frac{(b-a-1)}{3}\,\frac{(2a-k+1)^2(2a-k+2)}{2ak(2k-2a-2)}$$

$$\le \frac{(2a-k+1)(2a-k+2)}{3k}\left(-1 + \frac{(b-a-1)(2a-k+1)}{4a(k-a-1)}\right)$$

$$\le 0,$$

since $2a-k+1 \ge 1$, $k-a-1 \ge b-a-1$ and $2a > 2a-k+1$. Therefore $h \le 1$, proving the lemma. □

We are now ready to proceed in the evaluation of (1). By lemma 2,

$$P^{(d)}{}_{a_1,b_1}(k) \le \left[\frac{a_1^{b_1}\ a_2^{b_2}(n-b_2)^{n-a_1}(n-b_1)^{n-a_2}}{n^{2n+1}} \right]^d \qquad a_1,\, a_2 > 0.$$

Applying lemma 3(c) repeatedly, we obtain the following set of inequalities parameterized by i.

$$P^{(d)}{}_{a_1,b_1}(k) \le \left(\frac{d}{n}\right)^{d\,(d-i)} \left[\frac{a_1^{b_1}\ a_2^{b_2}(n-b_2)^{n-a_1}(n-b_1)^{n-a_2}}{n^{2n+1}} \right]^i \quad a_1,\, a_2 > 0 \qquad i = 0, 1, \ldots, d. \quad (3)$$

The following bound holds for all a_1, a_2. If $a_1 = 0$ or $a_2 = 0$ it follows from lemmas 2 and 3(d). Otherwise it follows from lemma 3(a)(b)(c).

$$P^{(d)}{}_{a_1,b_1}(k) \le \left(\frac{d}{n}\right)^{d(d-i)} \left(\frac{k}{n}\right)^{ik} \left(\frac{n-k}{n}\right)^{i(n-k)} \qquad i = 0, 1, \ldots, d. \qquad (4)$$

We are now able to derive upper bounds for the expected number of blocking k-pairs. We divide the proof into two parts, depending on whether k is less than or greater than $\frac{n}{2}$.

Lemma 5.

$$\textit{If} \quad 5 \le d \le k \le \frac{n}{2} \quad \textit{then} \quad \beta^{(d)}(k) \le \frac{1}{2\pi k} \left(\frac{d}{n}\right)^{d\,(d-4)}.$$

Proof. From (1) and (4) with $i = 4$ we obtain

$$\beta^{(d)}(k) \left(\frac{n}{d}\right)^{d\,(d-4)} \le \left(\frac{k}{n}\right)^{4k} \left(\frac{n-k}{n}\right)^{4\,(n-k)} \sum_{b_1+b_2=k} \sum_{a_1+a_2=k-1} \binom{n}{a_1}\binom{n}{a_2}\binom{n}{b_1}\binom{n}{b_2}.$$

Now an elementary calculation shows that

$$\sum_{a_1+a_2=k-1} \binom{n}{a_1}\binom{n}{a_2} \le \binom{2n}{k-1}$$

and

$$\sum_{b_1+b_2=k} \binom{n}{b_1}\binom{n}{b_2} \le \binom{2n}{k}.$$

Therefore,

$$\beta^{(d)}(k) \left(\frac{n}{d}\right)^{d\,(d-4)} \le \left(\frac{k}{n}\right)^{4k} \left(\frac{n-k}{n}\right)^{4\,(n-k)} \binom{2n}{k}^2 \quad \frac{k}{2n-k+1}$$

$$\leq \left(\frac{k}{n}\right)^{4k} \left(\frac{n-k}{n}\right)^{4\,(n-k)} \left[\frac{1}{\sqrt{2\pi k}} \left(\frac{2en}{k}\right)^{k}\right]^{2} \tag{5}$$

$$\leq \left(\frac{k}{n}\right)^{2k} \left(1-\frac{k}{n}\right)^{4(n-\frac{n}{2})} \frac{(2e)^{\,2k}}{2\pi k}$$

$$\leq 2^{-2k}\, e^{-2k}\, \frac{(2e)^{2k}}{2\pi k}$$

$$= \frac{1}{2\pi k}\,.$$

Inequality (5) follows from lemma 1(a). This completes the proof of lemma 5. □

Lemma 6

$$\mathit{If}\quad 5 \leq d \leq \frac{n}{2} \leq k \leq n \quad \mathit{then} \quad \beta^{(d)}(k) \;\leq \left(\frac{k}{e}\right)^{2} \left(\frac{d}{n}\right)^{d(d-2)}.$$

Proof. At first we assume that $a_1 > 0,\, a_2 > 0$ and $0 < b_1 \leq b_2 < n$. Let

$$f = \frac{a_1^{b_1} a_2^{b_2} (n-b_2)^{(n-a_1)} (n-b_1)^{(n-a_2)}}{b_1^{b_1} b_2^{b_2} (n-a_1)^{(n-a_1)} (n-a_2)^{(n-a_2)}}$$

and

$$g = \frac{a_1^{b_1} a_2^{b_2} (n-b_2)^{(n-a_1)} (n-b_1)^{(n-a_2)}}{a_1^{a_1} a_2^{a_2} (n-b_2)^{(n-b_2)} (n-b_1)^{(n-b_1)}}\; \frac{1}{n^2}\,. \tag{6}$$

Applying lemma 1(b) and (3) with $i = 2$, we see that

$$\binom{n}{a_1}\binom{n}{a_2}\binom{n}{b_1}\binom{n}{b_2} P^{(d)}{}_{a_1,b_1}\,(k) \;\leq \left(\frac{d}{n}\right)^{d(d-2)} fg.$$

We will show that $f \leq e^{-2}$ and $g \leq 1$.

The bound for f is obtained by applying the inequality $(1 + y/x)^x \leq e^y$ to the expression. Therefore

$$f = (1-\frac{b_1-a_1}{b_1})^{b_1}\,(1-\frac{b_2-a_2}{b_2})^{b_2}\,(1-\frac{b_2-a_1}{n-a_1})^{n-a_1}\,(1-\frac{b_1-a_2}{n-a_2})^{n-a_2}$$

$$\leq e^{-2}.$$

as $b_1 + b_2 = k$ and $a_1 + a_2 = k - 1$.

The bound for g is more lengthy. We begin by eliminating two variables from (6). Set $a_2 = a$, $a_1 = k - 1 - a$, $b_2 = b$, $b_1 = k - b$. Then we have

$$g = \left(\frac{k-1-a}{a}\right)^{a-b+1} \left(\frac{n-b}{n-k+b}\right)^{a+b-k} \frac{a(n-b)}{n^2} . \tag{7}$$

By our initial assumptions $d \le a \le k - d - 1$, $\frac{k}{2} \le b \le k < n$. We consider two main cases.

Case 1. $a \ge \frac{k-1}{2}$.

Since a and b are integers, $a + b \ge k$. Clearly

$$\frac{k-1-a}{a} \le 1 \qquad \text{and} \qquad \frac{n-b}{n-k+b} \le 1.$$

If $a \ge b$ then the first exponent of (7) is also greater than one and we are done. If $a = \frac{k-1}{2}$ we are also done because the first term is equal to one. The remaining case, $a \ge \frac{k}{2}$, $b > \frac{k}{2}$, is settled by lemma 4. Hence $g \le 1$.

Case 2. $a < \frac{k-1}{2}$.

In this case

$$\frac{k-1-a}{a} > 1, \quad \frac{n-b}{n-k+b} \le 1, \quad \text{and} \quad a - b + 1 \le 0.$$

If in addition, $a + b - k \ge 0$ or $b = \frac{k}{2}$ we are done immediately. Otherwise we have

$$a - b + 1 \le a + b - k \quad \text{and} \quad b - a \le k - 2a - 1.$$

Therefore, from (7) we obtain

$$g \le \left(\frac{k-1-a}{a} \; \frac{n-b}{n-k+b}\right)^{a+b-k} .$$

A simple calculation shows that

$$(k-1-a)\,(n-b) - a\,(n-k+b) = (k-2a-1)\,n + b - (b-a)\,k \ge 0.$$

Hence, $g \le 1$.

We now show that the initial assumptions can be lifted. We first observe that f and g are symmetric in the subscripts 1 and 2. Hence the above results hold if $b_1 > b_2$. Next, consider the degenerate case when $a_1 = 0$ and hence $a_2 = k - 1,\ b_1 = 0,\ b_2 = k$. The case $a_2 = 0$ is identical. A simple argument based on lemma 1(b) and equation (4) with $i = 2$ shows that

$$\binom{n}{k-1}\binom{n}{k} P^{(d)}{}_{0,0}(k) \le \frac{1}{2}\left(\frac{d}{n}\right)^{d(d-2)} .$$

The lemma now follows from the definition of $\beta^{(d)}(k)$ in equation (2) and the observation that the double sum involves at most $k^2 - k$ terms. □

We may now prove the main result of this section.

Proof of Theorem 1. Suppose (A, B) is a blocking k –pair with $|B| = k > n$. Then $(S - B, T - A)$ is a blocking $2n - k + 1$ –pair in the graph $G^*{}_d = (T, S, E_d)$, and $2n - k + 1 \le n$. Therefore the expected number of blocking pairs is bounded by

$$\beta^{(d)} \le 2\left\{ \sum_{k=d+1}^{n/2} \beta^{(d)}(k) + \sum_{k=n/2}^{n} \beta^{(d)}(k) \right\}$$

$$\le 2\left\{ \sum_{k=d+1}^{n/2} \left(\frac{d}{n}\right)^{d(d-4)} \frac{1}{2\pi k} + \sum_{k=n/2}^{n} \frac{k^2}{e^2} \left(\frac{d}{n}\right)^{d(d-2)} \right\}$$

$$\le 2\left\{ \left(\frac{d}{n}\right)^{d(d-4)} \frac{n}{4\pi d} + \left(\frac{d}{n}\right)^{d(d-2)} \frac{n^3}{2e^2} \right\}$$

$$\le 2\left(\frac{d}{n}\right)^{d^2-4d-1} \left\{ \frac{1}{4\pi} + \left(\frac{d}{n}\right)^{2d+1} \frac{1}{e^2}\,\frac{n^3}{2} \right\}$$

$$\le 2\left(\frac{d}{n}\right)^{d^2-4d-1} \left(\frac{1}{4\pi} + \frac{1}{2e^2} \right).$$

This completes the proof of the theorem. □

4. Expected Value of the Heuristic Solution

In this section we derive an upper bound on the expected value of the heuristic solution to the assignment problem described in section 2. Suppose initially that G_d contains a perfect matching M_d. We estimate the expected weight of a randomly chosen edge from G_d. We first recall that the expected value of the r-th order statistic from n uniform independent random variables in $[0, 1]$ is $\frac{r}{n+1}$[5]. Therefore the d directed edges selected from each vertex by the heuristic have expected weights $\frac{1}{n+1}, \frac{2}{n+1}, \ldots, \frac{d}{n+1}$. An arbitrarily selected edge therefore has expected weight $\frac{d+1}{2(n+1)}$. If G_d has a perfect matching, then such a matching chosen at random without regard to edge weights will have expected weight $\frac{n(d+1)}{n+1}$. The minimum weight perfect matching in such a graph therefore has expected weight bounded above by this amount. Let Z_d be the cost of the matching found by the heuristic. Then we have

$$E(\, Z_d \mid G_d \; has\ a\ perfect\ matching\,) \le \frac{n(d+1)}{n+1}\,.$$

However, G_d may not have a perfect matching. In this case, we take the trivial upper bound of $2n$ for Z_d. Hence

$$\begin{aligned}
E(\, Z_d\,) \le\ & E(\, Z_d \mid G_d \; has\ a\ perfect\ matching\,) + \\
& E(\, Z_d \mid G_d \; has\ no\ perfect\ matching\,)\ P(G_d \; has\ no\ perfect\ matching) \\
\le\ & \frac{n(d+1)}{n+1} + \frac{2n}{3}\left(\frac{d}{n}\right)^{d^2-4d-1},
\end{aligned}$$

by theorem 1. Setting $d = 5$ we obtain

$$\begin{aligned}
E(\, Z_d\,) &\le \frac{6n}{n+1} + \frac{2n}{3}\left(\frac{5}{n}\right)^{4} \\
&\le 6 \qquad \text{if } n \ge 11.
\end{aligned}$$

Since G_d is a subgraph of G_{d+1}, any matching in G_d exists also in G_{d+1}. Therefore,

$$E(\, Z_{d+1}) \le E(Z_{d)}.$$

Thus we have proved the following theorem.

Theorem 2.

$$\textit{If} \quad n \geq 11, \text{ and } d \geq 5 \ \textit{then} \quad E(\, Z_d \,) \leq 6.$$

We remark that the bound holds for $d = 5$ if the heuristic is modified to simply find any maximum cardinality matching in step 2, rather than one of minimum weight. Such a heuristic is considerably simpler to implement.

5. Acknowledgements

This research was supported by a grant from the Natural Science and Engineering Research Council under grant number A3013. We would like to thank an anonymous referee for a careful reading resulting, we hope, in many clarifications.

References

1. D. Avis, "Two Greedy Heuristics for the Weighted Matching Problem," *Proc. 9th S.E. Conf. on Comb., Graph Theory, and Computing*, pp. 65-76, Congressus Numerantium XXI, 1978.

2. D. Avis, "A Survey of Heuristics for the Weighted Matching Problem," *Networks*, vol. 13, pp. 475-493, 1983.

3. D. Avis and L. Devroye, "An Analysis of a Decomposition Heuristic for the Assignment Problem," *Operations Research Letters*, vol. 3, pp. 279-283, 1985.

4. A. Bondy and U.S.R. Murty, *Graph Theory with Applications,* American Elsevier, 1976.

5. H.A. David, *Order Statistics,* John Wiley, N.Y., 1970.

6. W.E. Donath, "Algorithms and Average Value Bounds for Assignment Problems," *IBM J. Research Dev.*, vol. 8, pp. 380-386, 1969.

7. R.M. Karp, "An Algorithm to Solve the m x n Assignment Problem in Expected Time O(mn log n)," *Networks*, vol. 10, pp. 143-152, 1980.

8. J.M. Kurtzberg, "On Approximate Methods for the Assignment Problem," *JACM*, vol. 9, pp.

419-439, 1962.

9. C.W. Lai, *A Heuristic for the Assignment Problem and Related Bounds,* SOCS 81.21, School of Computer Science, McGill University, 1981.

10. E.L. Lawler, *Combinatorial Optimization: Networks and Matroids,* Holt, Rinehart and Winston, N.Y., 1976.

11. D.S. Mitrinovic, *Analytic Inequalities,* Springer-Verlag, Heidelberg, 1970.

12. D. W. Walkup, ‘‘On the Expected Value of a Random Assignment Problem,’’ *SIAM J. Computing*, vol. 7, pp. 440-442, 1979.

13. D.W. Walkup, ‘‘Matching in Random Regular Bipartite Graphs,’’ *Discrete Mathematics*, vol. 31, pp. 59-64, 1980.