COMBINATORICA
Akadémiai Kiadó – Springer-Verlag

COMBINATORICA 16 (4) (1996) 453–464

DISTRIBUTED NEAR-OPTIMAL MATCHING

XIAOTIE DENG

Received June 20, 1995

In this paper, we consider the following distributed bipartite matching problem: Let $G = (L,R;E)$ be a bipartite graph in which boys are part L (left nodes), and girls are part R (right nodes.) There is an edge $(l_i,r_j) \in E$ iff boy l_i is interested in girl r_j. Every boy l_i will propose to a girl r_j among all those he is interested in, i.e., his adjacent right nodes in the bipartite graph G. If several boys propose to the same girl, only one of them will be accepted by the girl. We assume that none of the boys communicate with others. This matching problem is typical of distributed computing under incomplete information: Each boy only knows his own preference but none of the others. We first study the one-round matching problem: each boy proposes to at most one girl, so that the total number of girls receiving a proposal is maximized. If the maximum matching is M, then no deterministic algorithm can produce a matching of size not bounded by a constant, but a randomized algorithm achieves $\sqrt{M}$ — and this is shown optimal by an adversary argument. If we allow many rounds in matching, with the senders learning from their failures, then, for deterministic algorithms, the ratio (of the optimal solution to the solution of the algorithm) is unbounded when the number of rounds is smaller than $\Delta(G)$, and becomes bounded (two) at the $\Delta(G)$-th round. In contrast, an extension of the one-round randomized algorithm establishes that there is no such discontinuity in the randomized case. This randomized result is also matched by an upper bound of asymptotically the same order.

1. Introduction

We discuss a novel bipartite matching problem, motivated by recent interests in computational problems under incomplete information (see, e.g., [12, 9, 8, 7], for *online* problems involving in unknown future events, and [4,10,6] for *distributed optimization* problems). The competitive analysis [9], comparing a solution obtained under partial information with the optimal solution under complete information for the same set of input data, is a common methodology applied in these problems. Karp, et al., consider a centralized online matching problem while boys arrive in a linear order and each has to make a decision whom to match on arrival [8]. Each boy knows all the information available before his arrival. The unknown information is preferences of boys to arrive later.

For the *distributed matching* problem in our study, each boy knows his own information but not others. He may coordinate his strategy with those of the other boys but only as a function of his own information. Because of this restriction, any

Mathematics Subject Classification (1991): 68 Q 22, 68 R 10, 05 C 70 deterministic algorithm involving one round of proposals may end up with a matching which is constant in size — and arbitrarily smaller than the optimum matching size M (this is a counting argument spelled out in the proof of Theorem 2.1). However, we show that a simple randomized algorithm matches at least $\Omega(\sqrt{M})$ pairs — both in expectation and with high probability (Theorem 2.2). Furthermore, we show that $\sqrt{M}$ is the best possible performance by any randomized algorithm (Theorem 2.4). Moreover, we are also interested to know how much improvement the boys can make by learning from their experience being rejected (or accepted), in a way similar to that of [2] for improvement of shortest pathes by learning through repetitions of the same operation. In Section 3, boys are allowed to have several rounds to propose to girls and they can apply their past experience in previous rounds to make their decisions whom to propose to in later rounds. Obviously this would improve the overall performance in later rounds. We show that any deterministic algorithm with up to $\Delta(G)-1$ rounds may end up with only $O(\Delta(G))$ matched pairs — again, arbitrarily far from the optimum (Theorem 3.1). In contrast, with $\Delta(G)$ rounds we can obtain a maximal matching — at least $M/2$ pairs. We then use the high-probability version of Theorem 2.4 to show that randomized algorithms present no such gap (Theorem 3.3). Using the *competitive ratio* (defined as the worst case ratio of optimal matching size to the matching obtained by algorithms with incomplete information for maximization problems [8]), our results can be summarized as follows (with M for the maximum matching size).

1. The deterministic competitive ratio is $\Theta(M)$.
2. The randomized competitive ratio is $\Theta(\sqrt{M})$.
3. The deterministic competitive ratio for the i-th round is $\Theta(\frac{M}{i})$ for $i<\Delta(G)$, while the deterministic competitive ratio for the $\Delta(G)$-th round is 2.
4. The randomized competitive ratio for the i-th round is $\Theta(\frac{\Delta(G)}{i})$ for $i\leq\Delta(G)$.

2. One round matching problem

This section covers the one-round distributed matching problem in a bipartite graph (L,R,E). The setup of our problem is quite similar to that of *Linear Programming without Matrix* [10], in which each agent knows some local information while collectively they have complete information of the underlying structure. In our problem, each boy on a left node $u\in L$ observes only those edges incident to u. Each boy tries to choose an edge without communicating with other boys. They may, however, coordinate their effort by designing strategies that may benefit each other. We first focus on deterministic strategies.

Theorem 2.1. *For any set of deterministic strategies there is a bipartite graph so that the optimal matching has size $\Omega(n)$ and the matching constructed is $O(1)$.*

Proof. Let the adversary always present two edges to node L_i leading to R_j and $R_{(i+j) \bmod n}$ with j to be chosen later. Let X_{ik} to be 1 if L_i chooses R_k in exactly one of the two possible cases (i.e., another edge to R_{k+i} or R_{k-i}). Let X_{ik} to be 2 if L_i chooses R_k in both cases. Otherwise $X_{ik}=0$. Since the adversary has n possible ways, under the above restriction, to present two edges to L_i, we have $\sum_{k=1}^{n} X_{ik} = n$ for each $i : 1 \le i \le n$. It follows that $\sum_{i=1}^{n}\sum_{k=1}^{n} X_{ik} = n^2$. Hence, there exists some k^* such that $\sum_{i=1}^{n} X_{ik^*} \ge n$. Let $S = \{i : X_{ik^*} > 0\}$. Then $|S| \ge \frac{n}{2}$ since $X_{ik^*} \le 2$. Notice that, for $i \in S$, L_i can be presented with two edges: one is (L_i, R_{k^*}) and another must be either (L_i, R_{k^*+i}) or (L_i, R_{k^*-i}). For $i \in S$, we denote by $E(i)$ the edge other than (L_i, R_{k^*}) which is presented to L_i. Let $R(S) = \{x \in R : \text{x is incident to } E(i) \text{ for some } i \in S\}$.

For each $x \in R(S)$, assume $x = R_l$. Since it can only be incident to at most two such edges (L_{l+k*}, R_l) and (L_{l-k^*}, R_l), $|R(S)| \ge \frac{|S|}{2}$. Thus, there is a matching of size at least $n/4$, while all left nodes in S will choose R_{k^*}. ∎

In contrast, we show much improvement can be gained with the help of randomized algorithms.

Theorem 2.2. *If each left node proposes to one of its edges, picked uniformly at random, then the resulting matching is* $\Omega(\sqrt{M_{opt}})$*, where* M_{opt} *is the maximum matching, both in expectation and with probability* $1-o(1)$.

Proof. We first focus on the special class of bipartite graphs with perfect matchings and extend our proof to all the bipartite graphs later.

Let x_{ij} denote the probability that boy L_i assigns to choose R_j ($x_{ij}=0$ if there is no edge from L_i to R_j.) The probability j is chosen is $1 - \prod_{i=1}^{n}(1-x_{ij})$. Thus, the expected size of the matching is $EX[M] = \sum_{i=1}^{n}\left[1 - \prod_{i=1}^{n}(1-x_{ij})\right]$.

Denote $K = \{j : 1 \le \sum_{i=1}^{n} x_{ij}\}$ and $k = |K|$. We evaluate $EX[M]$ by considering edges in K and not in K separately.

1. $j \in K$: Apply the inequality: $0 \le 1-c \le e^{-c}$, for all $c \le 1$, we have $\prod_{i=1}^{n}(1-x_{ij}) \le e^{-\sum_{i=1}^{n} x_{ij}}$. It follows that,

$$1 - \prod_{i=1}^{n}(1-x_{ij}) \ge 1 - e^{-\sum_{i=1}^{n} x_{ij}}.$$

Since $\sum_{i=1}^{n} x_{ij} \geq 1$ for $j \in K$, we have, $1 - \prod_{i=1}^{n}(1 - x_{ij}) \geq 1 - e^{-1}$, for each $j \in K$. Therefore,

$$\sum_{j \in K} \left[1 - \prod_{i=1}^{n}(1 - x_{ij}) \right] \geq (1 - e^{-1})k.$$

2. $j \notin K$: Since $0 \leq x_{ij} \leq 1$ for all $1 \leq i, j \leq n$, we have

$$\prod_{i=1}^{n}(1 - x_{ij}) \leq 1 - \sum_{i=1}^{n} x_{ij} + \sum_{1 \leq i < l \leq n} x_{ij} x_{lj}.$$

With the inequality,

$$\sum_{1 \leq i < l \leq n} x_{ij} x_{lj} \leq \frac{1}{2} \sum_{i=1}^{n} x_{ij} \times \sum_{l=1}^{n} x_{lj},$$

we have

$$\prod_{i=1}^{n}(1 - x_{ij}) \leq 1 - \sum_{i=1}^{n} x_{ij} + \frac{1}{2}(\sum_{i=1}^{n} x_{ij})^2.$$

Combining this with the condition $\sum_{i=1}^{n} x_{ij} < 1$ $(j \notin K)$, we obtain $\prod_{i=1}^{n}(1 - x_{ij}) \leq 1 - \frac{1}{2} \sum_{i=1}^{n} x_{ij}$, which reduces to

$$1 - \prod_{i=1}^{n}(1 - x_{ij}) \geq \frac{1}{2} \sum_{i=1}^{n} x_{ij}.$$

Sum over all $j \notin K$,

$$\sum_{j \notin K} \left[1 - \prod_{i=1}^{n}(1 - x_{ij}) \right] \geq \sum_{j \notin K} \frac{1}{2} \sum_{i=1}^{n} x_{ij} = \frac{1}{2} \sum_{i=1}^{n} \sum_{j \notin K} x_{ij}.$$

To evaluate the sum, we consider two subcases:

2.1. Those left nodes i with $\sum_{j \notin K} x_{ij} = 0$: Since we use UNIFORM algorithm, for each such left node, all the adjacent right nodes are in K. By our assumption that there is a perfect matching in the bipartite graph, the total number of such left nodes i is no more than $|K| = k$ (Hall's Theorem).

2.2. Those left nodes i with $\sum_{j \notin K} x_{ij} > 0$: From above, we know there are at least $n-k$ such left nodes. Since UNIFORM takes the uniform distribution over all the edges, for each of these left nodes, $\sum_{j \notin K} x_{ij} \geq \frac{1}{k}$.

From these two subcases, we have

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j\notin K} x_{ij} \geq \frac{1}{2}(n-k)\frac{1}{k}.$$

Combining the above two cases, we have

$$\sum_{j=1}^{n}\left[1-\prod_{i=1}^{n}(1-x_{ij})\right] \geq (1-e^{-1})k+(n-k)\frac{1}{2k} = \Omega(\sqrt{n}).$$

In comparison, there is a perfect matching in the graph, the optimal solution is n.

In the general case there may not always be a perfect matching in bipartite graph, we need to apply the following corollary of Hall's Theorem (see, e.g., [3]).

Lemma 2.3. *Suppose the maximum matching of* $E=(U,V;E)$ *with* $|U|=|V|=n$ *is* m, *then for any* $S\subseteq U$, *we have* $|S|\leq|\Gamma(S)|+n-m$.

Thus, in the above proof, we only need to change Subcase 2.1 and 2.2. Let S be the set satisfying Subcase 2.1, i.e., $\Gamma(S)\subseteq K$. Then, we have $|S|\leq k+n-m$ and the number of left nodes i satisfying Subcase 2.2 is at least $n-(k+n-m)=m-k$. We have

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j\notin K} x_{ij} \geq \frac{1}{2}(m-k)\frac{1}{k}.$$

And hence we obtain a matching of expected size

$$\sum_{j=1}^{n}\left[1-\prod_{i=1}^{n}(1-x_{ij})\right] \geq (1-e^{-1})k+(m-k)\frac{1}{2k} = \Omega(\sqrt{m}).$$

We can show that this holds with probability $1-o(1)$ by applying the second momentum method [1, 11]. ∎

Furthermore, we can show that $\sqrt{M_{opt}}$ is the best one can achieve:

Theorem 2.4. *For any distributed randomized algorithm there is a bipartite graph so that the expected matching is* $O(\sqrt{M_{opt}})$, *where* M_{opt} *is the maximum matching size.*

Proof. We only consider the case where every left node is adjacent to a k-subset of right nodes. For each k-subset S of right nodes, the randomized strategy of an L_i assigns a nonnegative probability to each of the right nodes in S. We denote by $P_i(S)$ the right node assigned the smallest probability by L_i, breaking ties arbitrarily. Now consider all the $(k-1)$-subsets of right nodes: $\{T_j : 1\leq j\leq \binom{n}{k-1}\}$. For each such subset T_j, we order the right nodes not in T_j as $t_{j_0}, t_{j_1}, \cdots, t_{j_{n-k}}$

Let $v(1,(t_{j_s},T_j)) = 1$ if $S = \{t_{j_s}\} \cup T_j$ with $t_{j_s} = P_1(S)$, $1 \le s \le n-k+1$; zero otherwise. In general, let $v(i,(t_{j_s},T_j)) = 1$ if $S = \{t_{j_{(s+i-1) \bmod (n-k+1)}}\} \cup T_j$ with $t_{j_{(s+i-1) \bmod (n-k+1)}} = P_i(S)$; zero otherwise.

Since there are $\binom{n}{k}$ subsets S of size k, we have $\sum_{t_{j_s},T_j} v(i,(t_{j_s},T_j)) = \binom{n}{k}$ for every $i: 1 \le i \le n$. Therefore,

$$\sum_{i=1}^{n} \sum_{t_{j_s},T_j} v(i,(t_{j_s},T_j)) = n\binom{n}{k}.$$

Rewrite this as $\sum_{(t_{j_s},T_j)} \sum_{i=1}^{n} v(i,(t_{j_s},T_j)) = n\binom{n}{k}$. Since there are $n \times \binom{n-1}{k-1}$ possible choices of (t_{j_s},T_j), we know that there exists some T_{j^*} and s^* such that

$$\sum_{i=1}^{n} v(i,(t_{j^*_{s^*}},T_{j^*})) \ge \frac{n\binom{n}{k}}{n\binom{n-1}{k-1}} = \frac{n}{k}.$$

Let $I = \{i : v(i,(t_{j^*_{s^*}},T_{j^*})) = 1\}$. We know that $|I| \ge \frac{n}{k}$. By the definition of $v(i,(t_{j_s},T_j))$, when the boy L_i, $i \in I$, is presented with the subset $U_i = T_{j^*} \cup \{t_{j^*_{(s^*+i-1) \bmod (n-k+1)}}\}$ of right nodes, the probability of choosing $t_{j^*_{(s^*+i-1) \bmod (n-k+1)}}$ is at most $\frac{1}{k}$ and the probability of choosing an edge in T_{j^*} is at least $\frac{k-1}{k}$. The expected size of the matching that boys in $\{L_i : i \in I\}$ collectively construct would be at most $(k-1) + \frac{|I|}{k}$, where the first term corresponds to right nodes in T_{j^*}, and the second term corresponds to right nodes in $\cup_{i \in T}\{t_{j^*_{(s^*+i-1) \bmod (n-k+1)}}\}$. It is not difficult to see that the maximum matching is in the bipartite graph with the above edges is $|I|$. The bound follows now by choosing a size k^2 subset of I, which is possible when we choose $n = \Omega(k^3)$. ∎

3. Many rounds

In this section, we discuss how much information on the history of previous hits-and-misses can help when the unmatched senders are allowed to try again for more rounds.

Theorem 3.1. *Consider any two integers $m > \Delta > 0$. For any deterministic protocol, there exists a bipartite graph with the maximum degree Δ and the maximum matching size m such that in the k-th round, only a matching of size at most k is found for all $k < \Delta$, and a maximal matching of $\lceil \frac{m}{2} \rceil$ is found in the Δ-th round. These bounds can be achieved.*

Proof. To achieve these bounds, consider the following strategies: Each left node orders all its incident edges (arbitrarily). Propose to the corresponding right nodes according to the order, until a *hit* is observed (i.e., the proposal is accepted). After that, always make the same proposal. We call such type of strategies *hit-and-stay*. Let M_i be the matching obtained after the i-th round by the above strategy. We prove, by induction on i, that if M_{i-1} is not a maximal matching, then M_i has size at least i. First, the claim is trivially true for $i=1$ by using the convention $M_0=\emptyset$. Now assume it is true for i, now we want to prove it is true for the $i+1$-st round. Suppose now M_i is not a maximal matching in the bipartite graph. If $|M_i|>i$, then $|M_{i+1}|\geq i+1$ since $M_i \subseteq M_{i+1}$ according to our strategies of keeping an edge if we see a *hit*. Now assume $|M_i|=i$. Let $L(M_i)$ be the set of all left nodes incident to edges in M_i and $R(M_i)$ be right nodes. Let $S=L-L(M_i)$. Because of the hit-and-stay strategy, all previous choices of S fall in $R(M_i)$. If M_i is not a maximal matching, then at least one node u in S is connected to one right node not in $R(M_i)$. The node u would have a degree $>i$. Since all the previous i proposals of u are in $R(M_i)$ and $|R(M_i)|=i$, the $(i+1)$-st proposal of u will not in $R(M_i)$. Thus, $|M_{i+1}|$ has at least one edge other than those in M_i.

If we restrict our strategies to *hit-and-stay*, the upper bound, on the maximal obtainable matching in the worst case, can be shown by applying the method for the randomized lower bound in the previous section. The only change we need to make is to mark the last edge chosen by each boy in place of the edge with the least probability assignment. However, general strategies may require some boys to try other possibilities even after hits. The "last " chosen edge will depend on the past history of hits-and-misses and thus will not be well defined. Methods for Theorem 3.1 will not be applicable either, since strategies may depends on the names of right nodes a left node is connected to. To recover from these difficulties, we combine these two approach by defining "equivalent" strategies. Then, we apply Ramsey Theory to focus on one equivalent class of strategies.

Let $k=\Delta<n$. First we sort all the right nodes $R_1<R_2<\cdots<R_n$. Consider a strategy α of a left node when it is connected to $R_{i_1}<R_{i_2}<\cdots<R_{i_k}$. Let π be the order of index j such that R_{i_j} is chosen by α after a sequence of $k-1$ misses. Thus, for a strategy of the "equivalent" class, the order of right nodes being chosen after successive misses only depends on the relative orders of $R_{i_1}<R_{i_2}<\cdots<R_{i_k}$. We will say α is in the equivalent class π. Consider a set T of $(N=n^{\frac{1}{k}})$ right nodes, Assign each such k-subset of T to one left node in a 1-to-1 mapping. Give each k-subset the corresponding equivalent class number of the strategy by the left node assigned to that subset. Since there are $k=\Delta$ choice in each round, and we only consider $r\leq\Delta$ rounds, the total number of equivalent classes is no more than $r=\Delta^{\Delta}$. Now we need a Ramsey-type result on hypergraphs ([5], also see a proof in the appendix):

Lemma 3.2. *For any set of parameters $m, r, k > 0$ and $m, r \geq k$, consider a hypergraph $G = (T, E)$ with all edges in E having size k. Suppose each edge in G is marked by one of r numbers. If $|T| \geq k + A(k, mr)$, (where $A(\cdot,\cdot)$ is the Ackermann's function) then no matter how we number all size-k subsets of T, there is a size-m subset P of T such that all size-k subsets of P are marked with the same number.*

Applying the above Ramsey Lemma, we know that, for $m \geq k > 0$, if n is sufficiently large, there is a subset $R(0)$ of size $m + k - 1$ in T such that all the k-subsets of $R(0)$ have the same strategy number t^* and, without loss of generality, $t^* = (1, 2, \cdots, k)$ (i.e., on a sequence of up to k misses, the left node will choose right nodes in the increasing order of their names). For simplicity in notations, let $R(0) = \{R_1, R_2, \cdots, R_{m+k-1}\}$. Now we consider m subsets of $R(0)$: $\{R_1, R_2, \cdots, R_{k-1}\} \cup \{R_i\}$ assigned to left node B_i, $i = k, k+1, \cdots, m+k-1$. Notice that each left node B_i has a strategy of color $(1, 2, \cdots, k)$, $i = k, k+1, \cdots, m+k$. We will only present these subsets to corresponding m left nodes and all the other left nodes are presented nothing. The size of the maximum matching is m. In the first round, all the left node would choose 1 with one hit. The $m-1$ missed left nodes would make the same choice in the next round. In general before the i-th round, there are a total of $m - i + 1$ left nodes which has seen a sequence of $i - 1$ misses. They will make the same choice in the i-th round since they have the same history of $i - 1$ misses on the same order of right nodes, and their strategies are the same when restricted to a sequence of misses. Therefore, there are at most i hits in the i-th round for $i = 1, 2, \cdots, k-1$. We need a slightly different construction to obtain the upper bound of $\lceil \frac{m}{2} \rceil$ for the Δ-th round. ∎

Notice that there is a gap between round $\Delta(G) - 1$ and round $\Delta(G)$ in maximum sizes of matchings obtained by a deterministic algorithm. This gap of best achievable solutions disappears when randomized algorithms are allowed as shown by the following theorem. Notice this is also matched by an upper bound, asymptotically the same.

Theorem 3.3. *There is a randomized algorithm which achieves a matching size of at least $\Omega(\frac{r}{\Delta(G)} M_{opt})$ in the r-th round: $0 \leq r \leq \Delta(G)$. For any randomized algorithm, any $\Delta > 0$, there is a bipartite graph of the maximum degree $\Delta(G) = \Delta$ and the maximum matching size M_{opt}, on which the randomized algorithm can achieve a matching of size at most $O(\frac{r}{\Delta} M_{opt})$ in the r-th round: $0 \leq r \leq \Delta$.*

Proof. To obtain a matching of size $\Omega(\frac{r}{\Delta(G)} M_{opt})$, we adopt a variation of the UNIFORM algorithm: Each boy L_i chooses a random permutation $\pi(L_i)$ of all its incident edges, proposes to the corresponding girls in the order of $\pi(L_i)$ until succeeds and stays choosing the same girl in all the later rounds. Consider r-rounds of plays. Our algorithms is equivalent to the following: Each boy chooses r (to be precise, the maximum of r and its degree) edges to propose to. Then follow a hit-and-stay strategy. In other words, we first obtain a random subgraph H for which each left node has degree at most r, then follow hit-and-stay strategy. Let

$m(H)$ be the maximum matching size of H, then a matching of size at least $\frac{m(H)}{2}$ is constructed under the hit-and-stay strategy. Thus, it is sufficient to show that $m(H)=\Omega(\frac{r}{\Delta(G)}M_{opt})$. Consider a maximum matching (size M_{opt}). In obtaining H, each edge of the maximum matching is chosen with probability at least $\frac{r}{\Delta(G)}$ independently. The result follows.

We now adapt the proof of Theorem 2.4 to show that any randomized algorithm can guarantee a matching of size no more than $O(\frac{r}{\Delta}M_{opt})$ for all the graphs of maximum degree Δ. For each k-subset S of right nodes, in the first round, the randomized strategy of a boy L_i $(1\le i\le n)$ assigns a nonnegative probability to each of right nodes in S. After that, depending on history of hits or misses, he may change his probability assignments. This leads to an enormous amount of strategies. To resolve this difficulty, we only consider the probability assignments after all misses in previous rounds. Thus, let $p_{ij}^{(t)}$ be the probability assigned to girl R_j by boy L_i in the t-th round after $t-1$ misses, $1\le t\le r$.

We denote by $P_i(S)$ the right node j assigned the smallest accumulative probability $\sum_{t=1}^{r} p_{ij}^{(t)}$ by boy L_i, breaking ties arbitrarily. Thus $P_i(S)\le\frac{r}{k}$. Now consider all the $(k-1)$-subsets of right nodes: $\{T_j : 1\le j\le\binom{n}{k-1}\}$. For each such subset T_j, we order right nodes not in T_j as t_{j_0}, t_{j_1}, $\cdots$, $t_{j_{n-k}}$. Let $v(1,(t_{j_s},T_j))=1$ if $S=\{t_{j_s}\}\cup T_j$ with $t_{j_s}=P_1(S)$, $1\le s\le n-k+1$; zero otherwise. In general, let $v(i,(t_{j_s},T_j))=1$ if $S=\{t_{j_{(s-i+1)\bmod (n-k+1)}}\}\cup T_j$ with $t_{j_{(s+i-1)\bmod (n-k+1)}}=P_i(S)$; zero otherwise.

Since there are $\binom{n}{k}$ subsets S of size k, we have $\sum_{t_{j_s},T_j} v(i,(t_{j_s},T_j))=\binom{n}{k}$ for every $i:1\le i\le n$. Therefore,

$$\sum_{i=1}^{n}\sum_{t_{j_s},T_j} v(i,(t_{j_s},T_j)) = n\binom{n}{k}.$$

Rewrite it as $\sum_{t_{j_s},T_j}\sum_{i=1}^{n} v(i,(t_{j_s},T_j))=n\binom{n}{k}$. Thus, we know that there exists some T_{j^*} and s^* such that

$$\sum_{i=1}^{n} v(i,(t_{j^*_{s^*}},T_{j^*})) \ge \frac{n\binom{n}{k}}{(n-k+1)\binom{n}{k-1}} = \frac{n}{k}.$$

With the notation $I=\{i: v(i,(t_{j^*_{s^*}},T_{j^*}))=1\}$, we have $|I|\ge\frac{n}{k}$.

By the definition of $v(i,(t_{j_s},T_j))$, when the boy L_i, $i\in I$, is presented with the subset $U_i=T_{j^*}\cup\{t_{j^*_{(s^*+i-1)\bmod (n-k+1)}}\}$ of edges, the total probability of the first time choosing $t_{j^*_{(s^*+i-1) \bmod (n-k+1)}}$ after consecutive misses is at most $\frac{r}{k}$. Thus the expected number of edges not in T_{j^*} but chosen by boy L_i, $i \in I$ in the first time after consecutive misses is at most $\frac{r}{k}|I|$. In each round, at most $k-1$ edges from T_{j^*} can be chosen. Thus, the total number of other edges are no more than $r(k-1)$. Therefore, the total number of chosen right nodes is at most $r(k+\frac{|I|}{k})$. Since $k=\Delta$, our claim follows by choosing a size k^2 subset of I for $n \geq k^3$. ∎

Acknowledgement. This research is partially supported by a research grant from the National Science and Engineering Research Council of Canada. Many thanks are due to Christos Papadimitriou for his help in establishing Theorem 3.1 and 3.3.

References

[1] N. Alon, and J. H. Spencer: *Probabilistic Method*, Viley-Interscience Publication, John Wiley & Sons, Inc. Toronto, 1992.

[2] A. Blum, and P. Chalasani: An Online Algorithm for Improving Performance in Navigation, *FOCS93*, (1993), 2–11.

[3] B. Bollobás: *Extremal Graph Theory*, Academic Press Inc. New York, 1978.

[4] X. Deng, and C. Papadimitriou: Competitive Distributed Decision-Making, *Proc. 12th IFIP Congress*, Madrid, September 1993. And also to appear in Algorithmica.

[5] R. L. Graham, B. L. Rothschild, *and* J. H. Spencer: *Ramsey Theory*, John Wiley & Sons, New York, 1980.

[6] S. Irani, *and* Y. Rabani: On the Value of Information in Coordination Games. *FOCS93*, (1993), 12–21.

[7] E. Koutsoupias, *and* C. H. Papadimitriou: On the k-server conjecture, *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing*, (1994) 507–511.

[8] R. Karp, U. V. Vazirani, *and* V.V. Vazirani: An Optimal Algorithm for Online Bipartite Matching, *Proceedings of the Twenty-Second Annual ACM Symposium on Theory of Computing*, (1990), 352–358.

[9] M. S. Manasse, L. A. McGeoch, *and* D.D. Sleator: Competitive algorithms for on-line problems, *Journal of Algorithms*, **11** (1990), 208–230.

[10] C. H. Papadimitriou, *and* M. Yannakakis: Linear Programming without the Matrix, *STOC*, **25** (1993), 121–129.

[11] P. Raghavan: *Lecture Notes on Randomized Algorithms*, IRM Research Report, 1990, RC15340(#68237)1/9/90.

[12] D. D. Sleator, *and* R. E. Tarjan: Amortized efficiency of list update and paging rules, *Communications of the ACM*, **28**:2 (1985), 202–208.

4. Appendix

Proof of Lemma 3.2. We define the following type of generalized Ramsey numbers:

1. if $S_i > k$ for all i, then

$$R((C_1,S_1,k),\cdots,(C_r,S_r,k)) = 1 + R((C_1,S_1',k-1),\cdots,(C_r,S_r',k-1)),$$

   where $S_i' = R((C_1,S_1,k),\cdots,(C_i,S_i-1,k),\cdots,(C_r,S_r,k))$.

2. if $r > 1$ and $S_i = k$ for some i, then,

$$R((C_1,S_1,k),\cdots,(C_r,S_r,k)) = \\ = R((C_1,S_1,k),\cdots,(C_{i-1},S_{i-1},k),(C_{i+1},S_{i+1},k),\cdots,(C_r,S_r,k)).$$

3. if $r = 1$, then $R((C_1,S_1,k),\cdots,(C_r,S_r,k)) = S_1$.

4. if $k = 1$, then $R((C_1,S_1,1),\cdots,(C_r,S_r,1)) = \sum_{i=1}^{r} S_i - r + 1$.

We prove, by induction on k, that, if

$$n \geq R((C_1,S_1,k),(C_2,S_2,k),\cdots,(C_r,S_r,k)),$$

then, the following holds for one of $i : 1 \leq i \leq r$.

There is a subset of size S_i such that all its k-subsets are of the color C_i.

1. The base case: From the above definitions, it is true when $k = 1$ by the pigeonhole principle.
2. Assume the claim is true for $k-1$, now we consider k. For the same k, we do induction on $\sum_{i=1}^{r} S_i$.

   2.1 $r = 1$: Again the claim is trivial.

   2.2 $r > 1$ and one of $S_i = k$: W.l.o.g., let $S_r = k$. Then there is either one k-subset has color C_r or none has. In the former case, done. In the latter case, we know all the k-subsets are colored with C_1, $\cdots$, C_{r-1} and $\sum_{i=1}^{r-1} S_i$ decreases by k. The claim follows from the inductive hypothesis.

   2.3 $r > 1$ and $S_i > k$ for all $i : 1 \leq i \leq k$. Choose one point X. All the k-subsets containing X will induces subsets of size $k-1$ in the original hypergraph. Denote this by M. By the outer loop inductive hypothesis, we have: for one of i: there is a subset of size S_i' such that all its $k-1$-subsets are of the color C_i. For this i, by the inner loop inductive hypothesis, when $S_i' = R((C_1,S_1,k),\cdots,(C_i,S_i-1,k),\cdots,(C_r,S_r,k))$, then, either for some $j \neq i$, there is a subset of size S_j such that all its k-subsets are of the color C_j (we are done), or there is a subset of size $S_i - 1$ such that all its k-subsets are of the color C_i (we are done after adding the node X).

Let $S = \sum_{i=1}^{r} S_i$ and

$$f(k,S) = \max_{\sum_{i=1}^{r} S_i = S} R((C_1,S_1,k),(C_2,S_2,k),\cdots,(C_r,S_r,k)).$$

Then $f(k,S) \leq 1 + f(k-1, f(k,S-1))$ and it is not difficult to check that $f(k,S)-k$ is bounded by the Ackermann's function $A(k,S)$. Therefore, $k + A\left(k, \sum_{i=1}^{r} S_i\right)$ is an upper bound of $R((C_1,S_1,k),(C_2,S_2,k),\cdots,(C_r,S_r,k))$. ∎

Xiaotie Deng

Dept of CS,
York University, Ontario,
Canada M3J 1P3
deng@cs.yorku.ca