Available online at www.sciencedirect.com
ScienceDirect
Information Processing Letters 105 (2008) 150–154
Information Processing Letters
www.elsevier.com/locate/ipl

Feedback arc set problem in bipartite tournaments

Sushmita Gupta

School of Computing Science, Simon Fraser University, Canada

Received 18 December 2006; received in revised form 2 August 2007; accepted 7 August 2007

Available online 1 September 2007

Communicated by K. Iwama

Abstract

In this paper we give ratio 4 deterministic and randomized approximation algorithms for the FEEDBACK ARC SET problem in bipartite tournaments. We also generalize these results to give a factor 4 deterministic approximation algorithm for FEEDBACK ARC SET problem in multipartite tournaments.

Keywords: Approximation algorithms; Feedback arc set; Bipartite tournaments

1. Introduction

This paper deals with the problem of approximating feedback arc set in bipartite or multipartite tournaments. A directed graph, $D = (V, A)$, is a *multipartite tournament* if D is a directed orientation of a complete k-partite graph. When $k = 1$, it is an orientation of complete undirected graph and is known as a *tournament*. For $k = 2$ it is called a *bipartite tournament*. In feedback set problems we are given a graph G and a collection of cycles in G denoted by $\mathcal{C}$, the task is to find a minimum size set of vertices or arcs that meet all cycles in $\mathcal{C}$. More precisely the decision version of the problem is defined as follows:

FEEDBACK VERTEX (ARC) SET PROBLEM. Given a directed graph $D = (V, A)$, and a positive integer k, is there a subset of at most k vertices (or arcs) whose removal results in an acyclic graph?

FEEDBACK VERTEX/ARC SET (FVS/FAS) problem is one of the well-known NP-complete problems and is known to be complete even for a special graph class like directed graphs with in-degree and out-degree at most 3. While the FVS problem was known to be NP-complete in tournaments [9], a NP-completeness proof for the FAS problem in tournaments eluded us until recently. In 2005, three different groups independently showed FAS to be NP-complete in tournaments [2–4]. All the NP-completeness proofs for FAS in tournaments are very different from the usual NP-completeness proofs in that they use pseudorandom objects. The FAS problem has recently been shown to be NP-complete even in bipartite tournaments [6].

There is a long history of approximation algorithms for the feedback arc set problem in general directed graphs. Leighton and Rao [7] designed an $O(\log^2 n)$-factor approximation algorithm for feedback vertex set. Later, Seymour in [8] gave an $O(\log \tau^* \log\log \tau^*)$ approximation algorithm for feedback arc set where τ^* is the size of a minimum feedback vertex set. In 2005, a factor 2.5 randomized approximation algorithm was

E-mail address: gupta@cs.sfu.ca.

0020-0190/$ – see front matter
doi:10.1016/j.ipl.2007.08.023

developed for FAS in tournaments [1]. Later deterministic approximation algorithms with factor 5 and 3 were developed in [5] and [10], respectively.

Here we generalize the approximability results developed for FAS problem in tournaments to bipartite and multipartite tournaments by developing factor 4 deterministic and randomized approximation algorithms. Our results are inspired by the results presented in [1,10].

FAS problem in directed graphs, $D = (V, A)$ can also be phrased as, finding an ordering of vertices (V) such that the arcs from higher indices to lower indices, called *backward arcs*, are minimized. Arcs from lower indices to higher indices are called *forward arcs*. To see this, note that a directed graph D has a *topological ordering* if and only if D is a directed acyclic graph. We will use this equivalent formulation of FAS problem in our algorithms and will always give an ordering of vertices as our solution. The feedback arc set, F, can be formed by taking all the backward arcs in this ordering. In proving the upper bound on the approximation factor of our algorithms we use as lower bound, solution obtained from linear programming (LP) formulation of equivalent problems.

The paper is organized as follows. In Section 2 we give randomized factor 4 approximation algorithm for FAS problem in bipartite tournaments. Our deterministic approximation algorithm for FAS in bipartite tournaments is presented in Section 3. In this section we also suggest a simple modification to deterministic approximation algorithm for FAS in bipartite tournaments such that it can be made to work for multipartite tournaments. Finally we conclude with some remarks in Section 4.

Given a graph $D = (V, A)$ or $T = (V, A)$, n represents the number of vertices, and m represents the number of arcs. For a subset $V' \subseteq V$, by $D[V']$ (or $T[V']$) we mean the subgraph of D (T) induced on V'.

2. Randomized algorithm for FAS

In this section we give a factor 4 randomized approximation algorithm for FAS in bipartite tournaments. Our algorithm crucially uses the fact that a bipartite tournament has directed cycles if and only if it has a directed 4-cycle. Our randomized algorithm randomly selects an arc and partitions the bipartite tournament into two smaller ones around this arc and then solves the problem on these two smaller bipartite tournaments recursively.

In any directed graph, the size of minimum feedback arc set (MFAS) is at least the number of maximum arc disjoint directed cycles in the directed graph

Rand-MFASBT($T = (V = (V_1, V_2), A)$)

Step 0: If T is a directed acyclic graph then run the topological sort algorithm on T and let X be the order of V returned by the algorithm. **return**(X).

Step 1: Randomly select an arc $(i, j) \in A$, where $i \in V_1$, $j \in V_2$.

Step 2: Form the 2 sets V_L and V_R as follows.

$$V_L = \{u \mid (u, i) \in A \text{ or } (u, j) \in A\}$$

and

$$V_R = \{v \mid (i, v) \in A \text{ or } (j, v) \in A\}.$$

(It is clear that $T[V_L]$ and $T[V_R]$ are bipartite subtournaments of T.)

Step 3:
return(Rand-MFASBT($T[V_L]$), i, j, Rand-MFASBT($T[V_R]$)).

Fig. 1. Randomized algorithm for feedback arc set problem in bipartite tournaments.

(MADC). Since maximum number of arc disjoint 4-cycles (MAD4C) is upper bounded by the size of MADC, we have the following:

$$|\mathrm{MAD4C}| \leqslant |\mathrm{MADC}| \leqslant |\mathrm{MFAS}|.$$

We use the bound on MAD4C as a lower bound on MFAS in our algorithm and will show that the feedback arc set outputted by the algorithm is at most 4 times the size of MAD4C. This will prove the desired upper bound on the size of solution returned by the algorithm. The detailed description of our algorithm is presented in Fig. 1.

Now we show that the algorithm Rand-MFASBT is a randomized factor 4 approximation algorithm for feedback arc set in bipartite tournaments.

Theorem 1. *Let $T = (V = (V_1, V_2), A)$ be a bipartite tournament. Then the expected size of the feedback arc set (or the number of backward arcs in the ordering) returned by the algorithm* Rand-MFASBT(T) *is at most* 4 *times the size of a minimum feedback arc set of T.*

Proof. Let OPT denote the optimal solution of feedback arc set problem for $T = (V_1 \cup V_2, A)$ and C_{PIV} denote the size of backward arcs or feedback arc set returned by Rand-MFASBT on T. To prove the theorem we show that:

$$E[C_{PIV}] \leqslant 4OPT.$$

Observe that an arc $(u, v) \in A$ becomes a backward arc if and only if $\exists(i, j) \in A$ such that (i, j, u, v) forms a directed 4-cycle in T and (i, j) was chosen as the pivot when all 4 were part of the same recursive call. Pivoting on (i, j) will then put v in V_L and u in V_R. Thus making (u, v) a backward arc.

Let R denote the set of directed 4-cycles in T. For a directed 4 cycle r, denote by A_r the event that one of the arcs of r is chosen as a pivot when all four are part of the same recursive call. Let probability of the event A_r happening be $\Pr[A_r] = p_r$.

Define a random variable $X_r = 1$ when A_r occurs and 0 otherwise. We get

$$E[C_{PIV}] = \sum_{r\in R} E[X_r] = \sum_{r\in R} p_r. \tag{1}$$

As observed earlier the size of the set of arc disjoint 4-cycles is a lower bound on OPT. This is also true 'fractionally'. The following claim states this formally.

Claim. *If $\{\beta_r\}$ is a system of nonnegative weights on 4-cycles in R such that*

$$\forall a \in A, \qquad \sum_{\{r \mid r\in R, a\in r\}} \beta_r \leqslant 1,$$

then $OPT \geqslant \sum_{r\in R} \beta_r$.

Proof. Consider the following Covering LP:

$$\min \sum_{a\in A} x_a,$$
$$\sum_{a\in r} x_a \geqslant 1 \quad \text{for } r \in R, \quad \text{and}$$
$$x_a \geqslant 0 \quad \forall a \in A.$$

This is a LP formulation of hitting all directed 4-cycles of T and hence a solution to it forms a lower bound for OPT which considers hitting all cycles. If we take dual of this linear programming formulation then we get following Packing LP:

$$\max \sum_{r\in R} y_r,$$
$$\sum_{\{r \mid r\in R, a\in r\}} y_r \leqslant 1, \quad \forall a \in A, \quad \text{and}$$
$$y_r \geqslant 0, \quad \forall r \in R.$$

$\{\beta_r\}$ is a feasible solution to this dual LP and so a lower bound for the Covering LP and hence a lower bound on the OPT. □

We find such a packing using the probabilities p_r. Let $r = (i, j, k, l)$ be a directed 4-cycle then

$$\Pr\bigl[a = (i, j) \text{ is a pivot of } r \mid A_r\bigr] = \frac{1}{4}.$$

This is because we choose every edge as a pivot with equal probability. Let $a' = (k, l)$ and $B_{a'}$ be the event that a' becomes the backward arc. If A_r has occurred then a' becomes the backward arc if and only if (i, j) was chosen pivot among $(i, j), (j, k), (k, l), (l, i)$. So this gives us

$$\Pr[B_{a'} \mid A_r] = \frac{1}{4}.$$

So $\forall r \in R$ and $a \in r$ we have

$$\Pr[B_a \wedge A_r] = \Pr[B_a | A_r] \cdot \Pr[A_r] = \frac{1}{4}\Pr[A_r] = \frac{1}{4} p_r.$$

For 2 different 4-cycles r_1 and r_2, the events $B_a \wedge A_{r_1}$ and $B_a \wedge A_{r_2}$ are disjoint. From the moment a becomes a backward arc for r_1 the endpoints of a become part of different recursive sets, so the other event $B_a \wedge A_{r_2}$ can never occur. Therefore for any $a \in A$,

$$\sum_{\{r \mid a\in r\}} \Pr\bigl[B_a \wedge A_r\bigr] \leqslant \max_{\{r \mid a\in r\}} \left(\frac{1}{4} p_r\right) \leqslant 1,$$

since only one of the events in the summation can occur.

So $\{\frac{1}{4} p_r\}_{r\in R}$ is a fractional packing of R and hence a lower bound on OPT. Using Eq. (1) and the claim we get,

$$\frac{1}{4} E[C_{PIV}] = \frac{1}{4}\sum_{r\in R} p_r \leqslant OPT.$$

So $E[C_{PIV}] \leqslant 4OPT$ is proved as desired. □

3. Deterministic algorithm for FAS

We now give a deterministic factor 4 approximation algorithm for feedback arc set problem in bipartite tournaments. Our new algorithm is basically a derandomized version of the algorithm presented in the previous section. Here we replace the randomized step of Rand-MFASBT by a deterministic step based on a solution of a linear programming formulation of an auxiliary problem associated with feedback arc set problem in bipartite tournaments.

Given a bipartite tournament T, we call the problem of hitting all directed four cycles of T as the 4-CYCLE HITTING (4CH) problem. Notice that any feedback arc set of a bipartite tournament is also a solution to the 4CH problem. Given a bipartite tournament $T(V, A)$, we associate the following simple integer linear programming formulation to the 4CH problem. Let x_a be a variable for an arc $a \in A$, then

Det-MFASBT$(T = (V = (V_1, V_2), A))$

Step 0: If T is a directed acyclic graph then run the topological sort algorithm on T and let X be the order of V returned by the sorting algorithm. **return**(X).

Step 1: Select an arc $q = (i, j) \in A$, such that

$$\frac{|R_q|}{\sum_{a \in R_q} x_a^*}$$

is minimized.

Step 2: Form the 2 sets V_L and V_R as follows.

$$V_L = \{u \mid (u, i) \in A \text{ or } (u, j) \in A\}, \quad \text{and}$$
$$V_R = \{v \mid (i, v) \in A \text{ or } (j, v) \in A\}.$$

Step 3:
return(Det-MFASBT$(T[V_L]), i, j,$
Det-MFASBT$(T[V_R])$).

Fig. 2. Deterministic algorithm for feedback arc set problem in bipartite tournaments.

$$\min \sum_{a \in A} x_a$$
$$\text{s.t.} \sum_{a \in r} x_a \geqslant 1 \quad \text{for all directed four cycles } r \text{ in } T,$$
$$x_a \in \{0, 1\}, \quad \forall a \in A.$$

For the relaxation of the above LP formulation for 4CH problem we allow $x_a \geqslant 0$ for all $a \in A$. Let x^* be an optimal solution of this relaxed LP. We will use x^* as a lower bound for minimum feedback arc set in the approximation factor analysis of the algorithm. Notice that when we select an arc (i, j) as pivot for partitioning in our algorithm then all the arcs going from V_R to V_L become backward and these remain the same until the end of the algorithm. Given an arc $q = (i, j)$ and a bipartite tournament T, we associate a set of backward arcs and call it

$$R_q = \{(k, l) \mid (i, j, k, l) \text{ is a directed 4-cycle in } T\}.$$

We choose an arc q as pivot such that the size of R_q is 'minimized' and that would minimize

$$\frac{|R_q|}{\sum_{a \in R_q} x_a^*}.$$

We present our deterministic algorithm with the above mentioned change in Fig. 2.

We will show that the algorithm Det-MFASBT is a deterministic factor 4 approximation algorithm for feedback arc set in bipartite tournaments.

Theorem 2. *Let $T = (V = (V_1, V_2), A)$ be a bipartite tournament. Then the size of the feedback arc set (or the number of backward arcs in the ordering) returned by the algorithm* Det-MFASBT(T) *is at most 4 times the size of a minimum feedback arc set of T.*

Proof. Let X be the ordering returned by the algorithm Det-MFASBT(T) and let x^* be the optimal solution of the relaxed LP of 4CH problem. Also let B be the set of backward arcs with respect to the ordering X and OPT be the number of backward arcs in the ordering of T minimizing the backward arcs (size of minimum feedback arc set). Notice that

$$\sum_{a \in A} x_a^* \leqslant OPT.$$

Let $y_a = 1$ if $a \in B$ and $y_a = 0$ otherwise. Hence $\sum_{a \in A} y_a = \sum_{a \in B} y_a = |B|$. Now we show that

$$|B| = \sum_{a \in B} y_a \leqslant 4 \sum_{a \in B} x_a^* \leqslant 4 \sum_{a \in A} x_a^* \leqslant 4OPT.$$

If we take any $V' \subseteq V$ and consider the relaxed LP formulation of 4CH problem for $T[V']$ then x^* restricted to the arcs in $V[T']$ is a feasible solution to this LP. Notice that an arc (x, y) becomes backward for the ordering X when we choose an arc (i, j) as pivot and y is in V_L and x is in V_R of the partition based on (i, j) and this arc remains backward for subsequent recursions on the subproblems associated with each partition. In order to bound the size of backward arcs we show that there always exists a pivot $q = (i, j)$ such that

$$|R_q| \leqslant 4 \sum_{a \in R_q} x_a^*. \tag{2}$$

Notice that once we show this we are done as follows: we partition the arcs in B based on the pivot which has made it backward, let these partition be $B_{q_1}, B_{q_2}, \ldots, B_{q_l}$ where q_t $(1 \leqslant t \leqslant l)$, are the arcs chosen as pivot in the algorithm. Then

$$|B| = \sum_{t=1}^{l} |B_{q_t}| \leqslant \sum_{t=1}^{l} 4\Bigg(\sum_{a \in R_{q_t} = B_{q_t}} x_a^*\Bigg)$$
$$= 4 \sum_{a \in B} x_a^* \leqslant 4OPT.$$

Now we just need to show that we can choose a pivot q such that $|R_q| \leqslant 4 \sum_{a \in R_q} x_a^*$. Let R be the set of all directed four cycles in T, that is

$$R = \{((i, j), (j, k), (k, l), (l, i)) \text{ a directed 4-cycle in } T\}.$$

Define $x^*(r) = \sum_{a \in r} x_a$ for any $r \in R$. So we get the following:

$$\sum_{a\in A}\sum_{a'\in R_a}\mathbf{1}=\sum_{r\in R}\sum_{a''\in r}\mathbf{1}=\sum_{r\in R}4,\quad\text{and}$$
$$\sum_{a\in A}\sum_{a'\in R_a}x^*_{a'}=\sum_{r\in R}\sum_{a''\in r}x^*_{a''}=\sum_{r\in R}x^*(r).$$

Notice that for every $r \in R$, we have a constraint in the LP formulation of 4CH problem that $\sum_{a\in r} x_a \geqslant 1$ and hence $x^*(r) \geqslant 1$ for all $r \in R$. So we have

$$\sum_{r\in R}4\leqslant 4\left(\sum_{r\in R}x^*(r)\right).$$

This implies the existence of a pivot satisfying Eq. (2). If not then

$$\sum_{r\in R}4=\sum_{a\in A}\sum_{a'\in R_a}\mathbf{1}>\sum_{a\in A}4\left(\sum_{a'\in R_a}x^*_{a'}\right)$$
$$=4\left(\sum_{r\in R}x^*(r)\right),$$

which is a contradiction. This completes the proof of the theorem and shows that Det-MFASBT(T) is indeed a deterministic factor 4 approximation algorithm for FAS problem in bipartite tournaments. □

The algorithm Det-MFASBT(T) can be generalized to multipartite tournaments by doing a simple modification. When we pick an arc $q = (i, j)$ as pivot then we treat the partition containing i as one partition and union of all other parts as another one, making it 'similar' to a bipartite tournament. With this modification we create two smaller instances of multipartite tournaments in our algorithm and then recurse separately. All the analysis done for bipartite tournaments can be now carried over for multipartite tournaments. Without going into the details we state the following theorem.

Theorem 3. *Let $T(V, A)$ be a multipartite tournament. Then the modified* Det-MFASBT(T) *is a factor* 4 *approximation algorithm for FAS problem in T.*

4. Conclusion

In this paper we generalized the known approximation algorithms for FAS problem in tournaments to bipartite and multipartite tournaments. We gave factor 4 randomized and deterministic approximation algorithms for FAS problem in bipartite tournaments. There are only a few directed graph classes for which constant factor approximation algorithms for FAS problem are known. Here we add multipartite tournaments to graph classes for which there exists a constant factor approximation algorithm.

Acknowledgements

I thank Dr Ramesh Krishnamurti for the course he taught on approximation algorithms. Couple of papers I read for that course led to this work.

References

[1] N. Ailon, M. Charikar, A. Newman, Aggregating inconsistent information: ranking and clustering, in: Proceedings of 37th Annual ACM Symposium on Theory of Computing (STOC), 2005, pp. 684–693.

[2] N. Alon, Ranking tournaments, SIAM Journal on Discrete Mathematics 20 (1) (2006) 137–142.

[3] P. Charbit, S. Thomassé, A. Yeo, The minimum feedback arc set problem is NP-hard for tournaments, Combinatorics, Probability and Computing 16 (1) (2007) 1–4.

[4] V. Contizer, Computing Slater rankings using similarities among candidates, Technical Report RC23748, IBM Thomas J Watson Research Centre, NY, 2005.

[5] D. Coppersmith, L. Fleischer, A. Rudra, Ordering by weighted number of wins gives a good ranking for weighted tournaments, in: Proceedings of 17th Annual ACM–SIAM Symposium on Discrete Algorithms (SODA), 2006, pp. 776–782.

[6] J. Guo, F. Hüffner, H. Moser, Feedback arc set in bipartite tournaments is NP-complete, Information Processing Letters (IPL) 102 (2–3) (2007) 62–65.

[7] T. Leighton, S. Rao, Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms, Journal of ACM 46 (6) (1999) 787–832.

[8] P.D. Seymour, Packing directed circuits fractionally, Combinatorica 15 (1995) 281–288.

[9] E. Speckenmeyer, On feedback problems in digraphs, in: Proceedings of the 15th International Workshop on Graph-Theoretic Concepts in Computer Science (WG'89), in: Lecture Notes in Computer Science, vol. 411, 1989.

[10] A. van Zuylen, Deterministic approximation algorithm for clustering problems, Technical Report 1431, School of Operation Research and Industrial Engineering, Cornell University.