Graph Decomposition and a Greedy Algorithm for Edge-disjoint Paths

Kasturi Varadarajan* Ganesh Venkataraman†

Abstract

Given a directed graph G = (V, E) with n vertices and a parameter $l \geq 1$, we present an algorithm that finds a cut (set of edges) of size $O((n^2/l^2)log^2(n/l))$ whose removal separates every pair of vertices (s,t) in G such that the minimum distance between s and t in G is at least l. This theorem implies a nearly tight analysis of the greedy algorithm for finding edge-disjoint paths in directed graphs, and gives the best known approximation factor for this problem in terms of the number of vertices.

1. Introduction

We introduce certain conventions and notations that will be used throughout this paper. For any graph $G = (V, E)$ and any subset $C \subseteq E$ of edges, let $G \backslash C$ denote the graph $G = (V, E - C)$. $(u \rightsquigarrow v)_G$ denotes that there is a path from the vertex u to the vertex v in the graph G; $\neg(u \rightsquigarrow v)_G$ denotes that there is no such path. Let $\alpha(G, C) = \{$(u,v) $\mid (u \rightsquigarrow v)_G$ and $\neg(u \rightsquigarrow v)_{G\backslash C}$ $\}$. $d_G(u, v)$ represents the shortest path distance between u and v in G; we will drop the subscript when the graph G being referred to is clear. All graphs that we deal with are directed with unit edge costs unless otherwise specified. The main result of this paper is the proof of the following theorem:

Theorem 1.1 *For a directed graph G=(V,E) with n vertices and a parameter $l \geq 1$, there exists a set $C \subseteq E$ of $O((n^2/l^2)log^2(n/l))$ edges such that for every pair $u, v \in V$ such that $(u \rightsquigarrow v)_G$ and $d_G(u, v) \geq l$, there is no path from u to v in $G \backslash C$.*

The above theorem gives a nearly tight bound on the greedy algorithm for Edge Disjoint paths (EDPs) on directed graphs. In this problem, we are given a directed graph and a set of pairs of vertices and the goal is to connect as many pairs as possible by paths with the constraint that no two paths share an edge. Chekuri and Khanna [1] study a greedy algorithm for this problem: among all unrouted pairs, pick the closest in the graph consisting of yet unused edges, and connect then using the shortest path in this graph. Plugging Theorem 1.1 into their analysis, we obtain that the greedy algorithm is an $O((nlogn)^{2/3})$-approximation. Chekuri and Khanna proved a weaker version of Theorem 1.1 with a bound of $O(n^4/l^4)$ on C, which implied that the greedy algorithm is an $O(n^{4/5})$ approximation. Chekuri and Khanna [1] give a class of instances for which the greedy algorithm is an $\Omega(n^{2/3})$-approximation and hence the bound we show is nearly tight. Our bound is also the best known approximation factor for the EDP problem in terms of the number of vertices. The approximability of this problem in terms of the number of edges is better understood [2, 5, 4]; see the discussion in [1]. Independent of our work, Hajiaghayi and Leighton [3] prove a weaker version of Theorem 1.1 with a bound of $O(n^3/l^3)$ on the size of C; using this, they obtain an $O(n^{3/4})$ approximation for the EDP problem.

*University of Iowa, kvaradar@cs.uiowa.edu
†University of Iowa, gvenkata@cs.uiowa.edu

2. Key Lemma

Lemma 2.1 *Let G=(V,E) be directed graph with n vertices and $l \geq \log n$. Let s and t be any two vertices in G such that $(s \leadsto t)_G$ and $d(s,t) \geq l$. There exists a subset $C \subseteq E$ and a constant $c > 0$ such that*

$$\frac{|C|}{|\alpha(G,C)|} \leq \frac{c}{l^2} log^2(\frac{n}{l})$$

To prove the lemma we proceed to construct a series of graphs starting from G. Let $L_i = \{v | d(s,v) = i\}$ denote the set of vertices at the i'th level in a BFS starting from s. Note that $L_0 = \{s\}$ and $L_0, \ldots, L_l$ are non-empty. We use a parameter ϵ which will be defined later. As of now, we can treat ϵ as a small number such that $1/\epsilon$ is an integer. For $0 \leq j \leq 2/\epsilon$, let G_j be the subgraph of G induced by the vertices in levels $L_{j\epsilon l/8}, \ldots, L_{l-j\epsilon l/8}$. We call $L_{j\epsilon l/8}$ the "first level" in G_j and $L_{l-j\epsilon l/8}$ the "last level" in G_j. Now, define a series of threshold values corresponding to every G_j. The threshold value t_j for G_j equals $(n/l)^{j\epsilon}$. Let C_j denote a minimum cut in G_j that separates every vertex in the first level of G_j from every vertex in its last level. By our definitions of G_j, the last graph $G_{2/\epsilon}$ has at least $l/2$ levels of G in it. This implies that the min-cut $C_{2/\epsilon}$ in $G_{2/\epsilon}$ has size $O(n^2/l^2)$ which is $O(t_{2/\epsilon})$. The min-cut C_0 in G_0 has size at least $t_0 = 1$. Therefore, there exists a y with $1 \leq y \leq 2/\epsilon$ such that $|C_{y-1}| \geq t_{y-1}$ and $|C_y| = O(t_y)$. Choose such a y and consider the cut C_y. We will show below that

$$\alpha(G, C_y) = \Omega(|C_{y-1}|\epsilon^2 l^2).$$

This implies that

$$\frac{|C_y|}{|\alpha(G,C_y)|} = O(\frac{|C_y|}{|C_{y-1}|\epsilon^2 l^2}) = O(\frac{t_y}{t_{y-1}\epsilon^2 l^2}) = \frac{1}{l^2} O(\frac{1}{\epsilon^2}(\frac{n}{l})^{\epsilon}).$$

Setting $\epsilon = 1/\lceil \log(n/l) \rceil$, we get:

$$\frac{|C_y|}{|\alpha(G,C_y)|} = O(\frac{1}{l^2} log^2(\frac{n}{l})).$$

What is left to be shown is that $\alpha(G, C_y) = \Omega(|C_{y-1}|\epsilon^2 l^2)$. We first prove the following claim.

Claim 2.2 *Suppose $L_{i-1}, L_i, L_k, L_{k+1} \subseteq G_{y-1}$ where $i < k$. Then the number of pairs (u,v) such that $u \in L_{i-1} \cup L_i$ and $v \in L_k \cup L_{k+1}$ and $(u \leadsto v)_{G_{y-1}}$ is at least $|C_{y-1}|$.*

Proof: It is easy to see from the max-flow min-cut theorem, the integrality of the optimal flow, and flow decomposition of the optimal flow that there are $|C_{y-1}|$ edge-disjoint paths in G_{y-1} each of which starts from a vertex in the first level of G_{y-1} and ends at a vertex in its last level. We will argue that these paths must induce at least $|C_{y-1}|$ pairs as in the statement of the lemma. Any such path must use an edge (w,x) where $w \in L_{i-1}$ and $x \in L_i$ followed by an edge (y,z) where $y \in L_k$ and $z \in L_{k+1}$. For each path we fix one such quadruple (w,x,y,z) and say that the path uses edges (w,x) and (y,z) and "connection" (x,y).

Fix a connection (x,y) that is used by some path. Let S_1 (resp. S_2) be the set of all vertices $w \in L_{i-1}$ (resp. $z \in L_{k+1}$) such that some path uses edge (w,x) (resp. (y,z)). Let P_1 (resp. P_2) be the set of all paths that use edge (w,x) (resp. (y,z)) for some $w \in S_1$ (resp. $z \in S_2$). Note that $|S_1| = |P_1|$ and $|S_2| = |P_2|$ because the paths are edge-disjoint. We "generate" the set of pairs

$$\{(w,y)|w \in S_1\} \cup \{(x,z)|z \in S_2\}.$$

It is easy to see that $(u \leadsto v)_{G_{y-1}}$ for each generated pair (u,v). We delete the set of paths $P_1 \cup P_2$ and the vertices x and y. Note that the vertices x or y do not figure in any quadruple for the remaining paths. We have deleted at most $|P_1| + |P_2|$ paths but have generated $|P_1| + |P_2|$ pairs. By inductively repeating the argument on the remaining paths, the claim follows. $\square$

From the claim, it is easy to see that the number of pairs (u,v) such that $(u \rightsquigarrow v)_{G_{y-1}}$, $u \in L_i$ for some $(y-1)\epsilon l/8 \le i < y\epsilon l/8$, and $v \in L_k$ for some $l - y\epsilon l/8 < k \le l-(y-1)\epsilon l/8$ is $\Omega(|C_{y-1}|\epsilon^2 l^2)$. It is also easy to check that each such pair is in $\alpha(G, C_y)$.

Remark 2.3 A more careful argument shows that the number of pairs in $\alpha(G, C_y)$ is $\Omega(|C_{y-1}|\epsilon l^2)$. This improves the bound in the Lemma to $(1/l^2)\log(n/l)$.

3. Main Theorem

This section uses Lemma 2.1 to prove Theorem 1.1.

If $l \le log(n)$, we can remove all the edges and still get the bound defined in the above theorem. So we only consider the case when $l \ge log(n)$. The following algorithm gives a cut that satisfies the property stated in the theorem.

Algorithm Min-cut

1. $G_1 = G$, $i \leftarrow 1$
2. While $\exists$ s and t such that $(s \rightsquigarrow t)_{G_i}$ and $d_{G_i}(s,t) \ge l$
   (a) find a set C_i of edges in G_i such that $\frac{|C_i|}{|\alpha(G,C_i)|} = O(\frac{1}{l^2} log^2(\frac{n}{l}))$
   (b) $G_{i+1} \leftarrow G_i \backslash C_i$
   (c) $i \leftarrow i+1$

   end while
3. Return $\cup C_i$

From the key lemma, it is evident that for any s and t in G_i such that $d(s,t) \ge l$, a cut of the nature described in step 2(a) in the algorithm can be found. Hence the total cut size C is given by:

$$C = \sum_i C_i = O(\frac{1}{l^2} log^2(\frac{n}{l})) \sum_i |\alpha(G, C_i)|$$

But $\sum_i |\alpha(G, C_i)| \le n^2$ since each pair (u,v) of vertices contributes to at most one of the $\alpha(G, C_i)$, therefore:

$$C = O(\frac{n^2}{l^2} log^2(\frac{n}{l}))$$

Remark 3.1 Following Remark 2.3, we can improve the bound on C to $O(\frac{n^2}{l^2} log(\frac{n}{l}))$. This implies that the greedy algorithm approximates the EDP problem by a factor of $O(n^{2/3}\log^{1/3} n)$. For the case of undirected graphs, we can show a bound of $O(n^2/l^2)$ on the size of C in Theorem 1.1.

References

[1] Chandra Chekuri and Sanjeev Khanna. Edge Disjoint Paths Revisited. *Symposium on Discrete Algorithms*, 2003.

[2] V. Guruswami, S. Khanna, B. Shepherd, R. Rajaraman and M. Yannakakis. Near-Optimal Hardness Results and Approximation Algorithms for Edge-Disjoint Paths and Related Problems. Proceedings of 31st STOC, pp. 19-28, 1999.

[3] M. Hajiaghayi and F. T. Leighton. On max-flow min-cut ratio for directed multicommodity flows. Unpublished manuscript.

[4] S. G. Kolliopoulos and C. Stein. Approximating disjoint-paths problem using greedy algorithms and Packing Integer Programs. Integer Programming and Combinatorial Optimization, 1998.

[5] A.Srinivasan. Improved approximations for edge-disjoint paths, unsplittable flow, and related routing problems. Proceedings of the 38th Symposium on Foundationsof Computer Science, pp. 416-425, 1997.