The Minimum-Cost Congestion of Single-Sink Unsplittable Flows

An Honours Thesis Proposal

Yuchong Pan*

October 19, 2020

Abstract

In a capacitated directed graph with a sink and a set of terminals, we would like to concurrently route every demand along a *single* path from its terminal to the sink without violating the edge capacities, while minimizing the cost of the flow. This problem is NP-complete. We consider the *minimum-cost congestion* problem, measured by a bicriteria approximation (α, β) which means that a feasible, unsplittable flow satisfying all demands exists with its cost at most the optimum cost for fractional flows multiplied by β, if all edge capacities are multiplied by α. The *cut condition* and the *no bottleneck assumption* are assumed.

Dinitz et al. [3] achieve a 2-approximation for the minimum congestion problem without edge costs, and this result is close to the best possible because there exists an instance in which any unsplittable flow has a congestion arbitrarily close to 2. For the cost version, Skutella [12] gives a bicriteria $(3, 1)$-approximation algorithm, and this approximation ratio is the best known result. Goemans conjectures that the basic result of Dinitz et al. [3] can be generalized to the problem with edge costs [12]. This problem has been open for twenty years.

We observe that several similar graph-theoretic structures, namely zigzag-shaped cycles, are frequently present in approximation algorithms for solving problems on unsplittable flows and other classes of network flows such as confluent flows and d-furcated flows. However, these structures do not naturally fit problems with edge costs because we only augment along a single direction of such "cycles." We hope to extend these zigzag-shaped cycle structures for problems with edge costs during the course of this research.

1 Introduction

We consider the *single-sink unsplittable flow problem.* Let $G = (V, E)$ be a directed graph with a sink $t \in V$ and k terminals s_i, each of which has demand d_i, for $i \in [k]$. Let $u : E \to \mathbb{R}^+$ be edge capacities and $c : E \to \mathbb{R}^+$ edge costs. For each $i \in [k]$, we would like to route d_i units of flow from s_i to t along a single path. We say that a routing for a terminal is an *unsplittable flow* if it flows along a single path. The single-sink unsplittable flow problem is NP-complete; there exist polynomial-time reductions from well-known NP-complete problems, e.g. the *partition problem* and the *bin packing* problem, to the single-sink unsplittable flow problem [7, 8]. Hence, we search for approximation algorithms. Kleinberg [7, 8] introduces several approximation problems on single-sink unsplittable flows:

> ***Minimum Congestion.*** What is the smallest $\alpha \geq 1$ such that if all capacities are multiplied by α, a feasible, unsplittable flow satisfying all demands exists?

*yuchong.pan@alumni.ubc.ca, University of British Columbia.

Maximum Routable Demands. Find $T \subseteq [k]$ of demands which can be routed by a feasible, unsplittable flow and which maximizes $\sum_{i \in T} d_i$.

Minimum Number of Rounds. How many rounds are necessary to route all demands by feasible, unsplittable flows?

We consider the minimum congestion problem with edge costs, an "unsplittable" version of the well-known minimum-cost flow problem. Let C be the minimum cost of a feasible, fractional flow satisfying all demands. We use a *bicriteria approximation* (α, β) to measure a solution to the single-sink unsplittable flow problem in which a feasible, unsplittable flow $f : E \to \mathbb{R}^+ \cup \{0\}$ satisfying all demands with cost $\sum_{e \in E} f(e)c(e) \leq \beta \cdot C$ exists if all capacities are multiplied by α.

Let $c_{min} = \min_{e \in E} c(e)$ be the minimum capacity over edges $e \in E$. Let $d_{max} = \max_{i \in [k]} d_i$ be the maximum demand over $i \in [k]$. We assume the *no bottleneck assumption*, which states that the maximum demand is at most the minimum capacity, i.e. $d_{max} \leq c_{min}$. In addition, we assume the *cut condition*, which states that the sum of capacities in the directed cut $(S, V \setminus S)$ is at least the total demand $\sum_{i \in [k]} d_i$ for all $S \subseteq V \setminus \{t\}$. Given an instance of the single-sink unsplittable flow problem, we can convert it to an equivalent instance with a single terminal s of demand $\sum_{i \in k} d_i$ by adding a new vertex s and edges from s to every terminal t_i with edge capacity d_i, for $i \in [k]$. Assuming the cut condition, the existence of a feasible, fractional flow satisfying the demand in the converted graph is guaranteed by the classical maximum-flow minimum-cut theorem [6]:

Theorem 1 (Ford and Fulkerson [6])**.** *Let G be a capacitated graph with a single terminal s of demand d. Then there exists a feasible, fractional st-flow in G if and only if the sum of capacities in the directed cut $(S, V \setminus S)$ is at least d for all $S \subseteq V \setminus \{t\}$.*

Many classes of network flows are closely related, and techniques for solving one problem sometimes naturally extend to other classes. In addition to unsplittable flows, our interest particularly lies in *bounded degree* or *degree-constrained flows*, that is, feasible flows whose support networks have bounded degrees at each vertex. We say that a flow is *d-furcated* if the out-degree of each vertex is at most d in the support network of the flow. In particular, the following special cases of d-furcated flows, possibly with additional constraints, are of our interest:

Fractional flows. Each flow can send flow *fractionally* to any number of outgoing edges. Hence, every demand can be routed *fractionally* along *any* path from a source to the sink.

Unsplittable flows. The routing for each terminal must flow along a *single* path.

Confluent flows. Each vertex can send flow to at most one outgoing edge. In other words, the support network of flow forms a directed arborescence rooted at the sink. Note that a confluent flow implies that the flow is unsplittable.

β-confluent flows. Each vertex must send at least a β-fraction of its total flow to a single outgoing edge, and all remaining flow to at most one other outgoing edge.

Halfluent flows. Each vertex must either send its total flow on one outgoing edge, or split its total flow *equally* to two outgoing edges.

Bifurcated flows. Each vertex can send flow to at most two outgoing edges.

2 Present Work

The minimum congestion problem and its cost version on single-sink unsplittable flows were first considered by Kleinberg [7, 8]. For the minimum congestion problem without edge costs, Kleinberg [7, 8] gives a 16-approximation algorithm for directed graphs; on undirected graphs, Kleinberg [7] obtains an approximation ratio of $\frac{9}{2} + \sqrt{14} \approx 8.25$. Kolliopoulos and Stein [9] improve the approximation ratio to $\frac{10}{3} \approx 3.33$ for both directed and undirected graphs; in a later publication of their work, the approximation ratio is further improved to 3 [10]. Dinitz et al. [3] obtain a 2-approximation algorithm using the technique of "alternating cycles", and this approximation ratio is close to the best possible because there exists an instance in which any unsplittable flow violates the capacity of some edge by a congestion arbitrarily close to 2. If the *no bottleneck assumption* is not assumed, Dinitz et al. [3] mention that a 5-approximation algorithm can be obtained, which improves the approximation ratio of $3 + 2\sqrt{2} \approx 5.8285$ by Kolliopoulos and Stein [9, 10].

For the minimum congestion problem with edge costs, Kleinberg [7, 8] achieves a bicriteria $(6 + 2\sqrt{5}, 3 + 2\sqrt{5})$-approximation on undirected graphs. Kolliopoulos and Stein [10] provide a bicriteria $(3, 2)$-approximation algorithm for both directed and undirected graphs. Asano [1] generalizes the technique of Kolliopoulos and Stein [10] to the minimum-cost congestion problem with k sinks, obtaining a bicriteria $(k + 2, 2)$-approximation. The best ratio for the minimum-cost congestion problem so far is the bicriteria $(3, 1)$-approximation with the *no bottleneck assumption* and the bicriteria $(3 + 2\sqrt{2}, 1)$-approximation with arbitrary demands, both provided by Skutella [12].

On the negative side, several inapproximability results are present. Lenstra et al. [11] prove that there is no polynomial-time algorithm that can achieve a worst-case ratio less than $\frac{3}{2}$ for the *scheduling problem with unrelated parallel matchines* unless P = NP. Since this problem can be reduced to the minimum congestion problem on single-sink unsplittable flows, it follows that the minimum congestion problem cannot be approximated with a ratio better than $\frac{3}{2}$ in polynomial time unless P = NP. Kolliopoulos and Stein [10] show that there is no ρ-approximation with $\rho < 2$ for the minimum congestion problem on unsplittable flows with two sinks unless P = NP. For the cost version of the minimum congestion problem, Erlebach and Hall [5] obtains an inapproximability result of the bicriteria $(2 - \varepsilon, 1)$-approximation for arbitrary $\varepsilon > 0$.

3 Techniques

A typical starting point throughout the literature to tackle the minimum congestion problem and its cost version on single-sink unsplittable flows and related classes of network flows is any feasible, fractional flow in the given network [e.g., 3, 4, 7, 8, 9, 10, 12]. The existence of a feasible, fractional flow is guaranteed by the *cut condition* and Theorem 1 [6]. In addition, Dinitz et al. [3] show that an algorithm that leads to a ν-approximation *with* the cut condition also implies the same approximation ratio if the cut condition is not satisfied. By solving a parametric network flow problem (e.g. by binary search), we can find the smallest $\alpha \geq 1$ such that the cut condition is satisfied by multiplying all edge capacities by α. Applying the same algorithm yields an unsplittable flow in which the total flow on any edge $e \in E$ is upper bounded by $\nu \cdot \alpha c(e)$. Since α is a lower bound on the optimum congestion for fractional flows, the algorithm induces a ν-approximation.

Several similar graph-theoretic structures, namely zigzag-shaped cycles, frequently appear in approximation algorithms for single-sink unsplittable flows and other related classes of network flows. For instance, an *alternating cycle* in [3] is a directed "cycle" structure formed by alternating between maximal *forward paths* and maximal *backward paths*, which are directed paths of opposite directions. Additionally, a backward path only consists of *singular* incoming edges whose reachable vertices form a directed path. Another zigzag-shaped cycle structure used in approximation algorithms for solving network flow problems is the sawtooth cycle structure. In the $(1 + \log k)$-congestion algorithm of Chen et al. [2] for confluent flows, a *sawtooth cycle* is a directed "cycle" of length > 2 which allows edges between sinks and their neighbours (i.e. *frontier* vertices) to have the reverse direction. More generally, a *(general) sawtooth cycle* in the $(1+\frac{1}{d-1})$-congestion algorithm of Donovan et al. [4] for d-furcated flows consists of a collection of paths $\{(u_1, v_1), P_1, \ldots, (u_r, v_r), P_r\}$, where $(u_i, v_i) \in E$ and P_i is a u_{i+1}-v_i directed path for all $i \in \{1, \ldots, r\}$ with subscripts modulo r.

Zigzag-shaped cycles are usually unwanted in the realm of unsplittable or degree-constrained network flows, because the alternation between directed paths of opposite directions often indicates the existence of vertices that receive flow from multiple paths, or send flow to multiple outgoing edges, at the intersections of the alternating directed paths. Therefore, a common operation used in approximation algorithms for unsplittable or degree-constrained network flows is to *eliminate* or *break* zigzag-shaped cycles by augmenting along these cycles. We also observe that such augmentations typically add ε flow to each backward (resp. short) paths and subtract ε flow to each forward (resp. long) paths, where the value of ε is chosen so that either the flow on an edge along the cycle vanishes, or another operation may proceed (e.g. moving terminals in [3]). Such augmentations preserve the total incoming flow and the total outgoing flow of each vertex, and guarantee that the new flow also satisfies the edge or vertex capacity constraints.

Furthermore, Donovan et al. [4] prove several graph-theoretic characterizations of the absence of (general) sawtooth cycles in a directed acyclic graph G, hence relating the sawtooth cycle structure, the bipartite *auxiliary directed grpah* D induced by G, and the *digon-tree representation* $\mathcal{D}$ induced by D. Specifically, an *auxiliary directed graph* D is constructed from a directed acyclic graph G by creating two bipartite copies v^+ and v^- with a *vertex edge* (v^-, v^+) for each vertex $v \in V$, and by adding a reverse edge for each edge $e \in E$. Moreover, a *digon-tree representation* $\mathcal{D}$ is obtained from a directed graph D by contracting the digon[1] edges of D, if the underlying undirected graph edge-induced by the digons of D form a forest. Donovan et al. [4] prove Theorem 2 which characterizes the absence of (general) sawtooth cycles in a directed acyclic graph G. With this characterization, Donovan et al. [4] prove Theorem 3 regarding vertex-induced trees in a directed acyclic graph G without sawtooth cycles. This theorem plays an essential role in the $(1 + \frac{1}{d-1})$-congestion algorithm of Donovan et al. [4] because it provides a layered structure of the original graph G on which the algorithm redirects flow layer by layer, guaranteeing that the congestion at the next layer is upper bounded by $1 + \frac{1}{1+d}$ in each iteration.

Theorem 2 (Donovan et al. [4])**.** *In a directed acyclic graph G, the following are equivalent:*

1. *G does not contain a (general) sawtooth cycle.*
2. *The auxiliary directed graph D induced by G does not contain a (simple) cycle of length at least three (and hence at least four since D is bipartite).*
3. *The auxiliary directed graph D induced by G has an acyclic digon-tree representation $\mathcal{D}$.*

Theorem 3 (Donovan et al. [4])**.** *Let $G = (V, E)$ be a directed acyclic graph that contains no sawtooth cycles. Then G is the union of a set $\mathcal{T}$ of vertex-induced trees in G such that any vertex $v \in V$ is contained in at most two of the trees. Moreover, for a vertex $v \in V$, all outgoing edges from v are contained in the same tree, and all incoming edges at v are contained in the same tree.*

[1] A *digon* is a pair of edges that form a simple directed cycle of length 2.

We conclude this section by commenting that the zigzag-shaped cycle strucutres are hard to use in the cost version of the minimum congestion problem on unsplittable or degree-constrained network flows. One of the reasons for the absence of this technique in the cost version is the single-direction augmentation along zigzag-shaped cycles. That is, we always add ε flow to each backward (resp. short) paths and subtract ε flow to each forward (resp. long) paths. The augmentation is not performed in the other direction, so the total flow on each edge or vertex of forward (resp. long) paths does not increase, hence preserving the capacity constraints. In the cost version, however, the flow along the other direction may have a lower cost, in which case the augmentation along the other direction *should* be preferend. It may be possible and interesting to extend current techniques for the existing zigzag-shaped cycle structures or to devise new zigzag-shaped cycle structures for solving the cost version on unsplittable or degree-constrained network flows.

4 Open Problems

Many interesting problems remain open in the realm of unsplittable and degree-constrained network flows. For instance, Dinitz et al. [3] achieve a congestion of 2 for single-sink unsplittable flows *without* edge costs. However, the best known result for the cost version is the bicriteria $(3, 1)$-approximation by Skutella [12]. We note that Skutella [12] also obtains a bicriteria $(2, 1)$-approximation for the cost version with *rounded* demands, i.e. $d_i \mid d_j$ or $d_j \mid d_i$ for each $i, j \in [k]$, whereas in the general case an additional factor of 2 is present in the algorithm. According to [12], Goemans conjectures in a personal communication in January 2000 that the basic result of Dinitz et al. [3] can be generalized to the minimum-cost congestion problem on single-sink unsplittable flows.

The current $(1 + \frac{1}{d-1})$-approximation of Donovan et al. [4] for the minimum congestion problem on d-furcated flows assumes unit vertex capacities. It is natural to ask whether this result continues to hold with general vertex capacities. As noted in Section 3, the zigzag-shaped cycle structures do not naturally fit problems with edge costs, and no results are known for the minimum-cost congestion problem on confluent flows or other d-furcated flows. For halfluent flows, it is unknown whether the minimum congestion is bounded by a constant. Donovan et al. [4] conjectures that the techniques of the sawtooth cycle structure may be extended to obtain an $O(1)$-congestion for halfluent flows. It may also be interesting to study open problems on degree-constrained flows in addition to unsplittable flows, as many techniques apply in both classes of network flows.

5 Timeline and Deliverables

The research on the minimum-cost congestion problem of single-sink unsplittable flows consists of two phases. Phase I of the research plan includes a literature review in the research area of the network flow theory, and primarily in the area of unsplittable and degree-constrained network flows. In particular, three papers, [3], [4] and [12], are to be digested carefully and thoroughly, and presented to the supervisor. As a result of the literature review, a tentative outline of the honours thesis is expected to be produced by the end of October 2020.

Phase II of the plan consists of research on the minimum-cost congestion problem of single-sink unsplittable flows, which involves weekly meetings and other discussions with the supervisor. A set of milestones to be completed during Phase II is listed as follows. The outline of the honours thesis is expected to be finalized by mid-January 2021. A survey on the connections with other relevant problems (e.g. unsplittable flows on rings, trees, etc.) and on the history of these problems will also be completed by mid-January 2021. Finally, a draft of the honours thesis is expected to be produced by mid-February 2021, and the honours thesis should be finalized, submitted and presented by the end of the April exam period of 2021.

References

[1] Y. Asano. Experimental evaluation of approximation algorithms for the minimum cost multiple-source unsplittable flow problem. In *ICALP Satellite Workshops*, pages 111–122, 2000.

[2] J. Chen, R. D. Kleinberg, L. Lovász, R. Rajaraman, R. Sundaram, and A. Vetta. (almost) tight bounds and existence theorems for confluent flows. In *Proceedings of the thirty-sixth annual ACM symposium on Theory of computing*, pages 529–538, 2004.

[3] Y. Dinitz, N. Garg, and M. X. Goemans. On the single-source unsplittable flow problem. *Combinatorica*, 19(1):17–41, 1999.

[4] P. Donovan, B. Shepherd, A. Vetta, and G. Wilfong. Degree-constrained network flows. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 681–688, 2007.

[5] T. Erlebach and A. Hall. Np-hardness of broadcast scheduling and inapproximability of single-source unsplittable min-cost flow. *Journal of Scheduling*, 7(3):223–241, 2004.

[6] L. R. Ford and D. R. Fulkerson. Maximal flow through a network. *Canadian journal of Mathematics*, 8:399–404, 1956.

[7] J. M. Kleinberg. *Approximation algorithms for disjoint paths problems*. PhD thesis, Massachusetts Institute of Technology, 1996.

[8] J. M. Kleinberg. Single-source unsplittable flow. In *Proceedings of 37th Conference on Foundations of Computer Science*, pages 68–77. IEEE, 1996.

[9] S. G. Kolliopoulos and C. Stein. Improved approximation algorithms for unsplittable flow problems. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 426–436. IEEE, 1997.

[10] S. G. Kolliopoulos and C. Stein. Approximation algorithms for single-source unsplittable flow. *SIAM Journal on Computing*, 31(3):919–946, 2001.

[11] J. K. Lenstra, D. B. Shmoys, and É. Tardos. Approximation algorithms for scheduling unrelated parallel machines. *Mathematical programming*, 46(1-3):259–271, 1990.

[12] M. Skutella. Approximating the single source unsplittable min-cost flow problem. *Mathematical Programming*, 91(3):493–514, 2002.