GRAPH PARTITIONING WITH DEMANDS: GENERALIZED GRAPH CONDUCTANCE WITH APPLICATIONS TO HIERARCHICAL CLUSTERING

Michał Szyfelbein and Dariusz Dereniowski

Gdańsk University of Technology

e-mail: michal.szyfelbein@pg.edu.pl, dariusz.dereniowski@pg.edu.pl

In this work, we study various graph partitioning tasks under a general demand model. In each such task, we are given a graph $G = (V, E, c, w)$ with a capacity function $c\colon E \to \mathbb{N}$ and a demand function $w\colon V \times V \to \mathbb{N}$. Our main focus is the problem of finding a cut $(S, \bar{S})$ minimizing the quantity

$$\psi_w(S) = \frac{c\big(S, \bar{S}\big)}{w(S, V) \cdot w\big(\bar{S}, V\big)}.$$

Here, $c\big(S, \bar{S}\big)$ is the cost of edges between S and the complement of S, $\bar{S}$, and $w(S, V) = w(S) + w\big(S, \bar{S}\big)$ is the sum of the internal demand within S, denoted $w(S)$, and the demand between vertices of S and $\bar{S}$, denoted $w\big(S, \bar{S}\big)$. We call $\psi_w(S)$ the *generalized conductance* of the cut $(S, \bar{S})$, and the task of minimizing $\psi_w(S)$ the Generalized Minimum Conductance Problem. Our main contribution is an algorithm with an $O(\log n)$-approximation guarantee for this objective. This result is achieved via a two-way reduction: first to the well-known generalized k-multicut problem, and then to a constrained variant of the classic sparsest-cut problem, with an additional upper-bound constraint on the amount of demand that may be cut.

Moreover, we show that this result can be leveraged to obtain an $O(\log n)$ bicriteria approximation for Graph Partitioning with Demands, where the goal is to find a minimum-cost subset of edges C such that for every component H of $G \setminus C$, $w(H) \leq w(V)/2$. This, in turn, yields an $O(\log n)$-approximation for Hierarchical Clustering with Demands, the problem of finding a hierarchy of cuts that partitions the graph into increasingly refined clusters.

References

[1] Sanjeev Arora, James R. Lee, and Assaf Naor. Euclidean distortion and the sparsest cut. In *Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '05, page 553–562, New York, NY, USA, 2005. Association for Computing Machinery.

[2] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56(2), April 2009.

[3] Moses Charikar and Vaggos Chatziafratis. Approximate hierarchical clustering via sparsest cut and spreading metrics. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '17, page 841–854, USA, 2017. Society for Industrial and Applied Mathematics.

[4] Vincent Cohen-addad, Varun Kanade, Frederik Mallmann-trenn, and Claire Mathieu. Hierarchical clustering: Objective functions and algorithms. *J. ACM*, 66(4), June 2019.

[5] Uriel Feige, MohammadTaghi Hajiaghayi, and James R. Lee. Improved approximation algorithms for minimum-weight vertex separators. In *Proceedings of the Thirty-Seventh Annual*

ACM Symposium on Theory of Computing, STOC '05, page 563–572, New York, NY, USA, 2005. Association for Computing Machinery.

[6] Jochen Könemann, Ojas Parekh, and Danny Segev. A unified approach to approximating partial covering problems. *Algorithmica*, 59(4):489–509, April 2011.

[7] Harald Räcke. Optimal hierarchical decompositions for congestion minimization in networks. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing*, STOC '08, page 255–264, New York, NY, USA, 2008. Association for Computing Machinery.