van Bevern, René; **Bredereck, Robert**; **Chopin, Morgan**; **Hartung, Sepp**; **Hüffner, Falk**; **Nichterlein, André**; **Suchý, Ondřej**

Fixed-parameter algorithms for DAG partitioning. (English) Zbl 1355.05204

Discrete Appl. Math. 220, 134-160 (2017).

Summary: Finding the origin of short phrases propagating through the web has been formalized by *J. Leskovec* et al. ["Meme-tracking and the dynamics of the news cycle", in: Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining, KDD'09, Paris, France, June 28 – July 01, 2009, New York, NY: Association for Computing Machinery (ACM). 497–506 (2009; doi:10.1145/1557019.1557077)] as DAG partitioning: given an arc-weighted directed acyclic graph on n vertices and m arcs, delete arcs with total weight at most k such that each resulting weakly-connected component contains exactly one sink – a vertex without outgoing arcs. DAG partitioning is NP-hard.

We show an algorithm to solve DAG partitioning in $O(2^k \cdot (n + m))$ time, that is, in linear time for fixed k. We complement it with linear-time executable data reduction rules. Our experiments show that, in combination, they can optimally solve DAG partitioning on simulated citation networks within five minutes for $k \leq 190$ and m being 10^7 and larger. We use our obtained optimal solutions to evaluate the solution quality of Leskovec et al.'s heuristic.

We show that Leskovec et al.'s heuristic works optimally on trees and generalize this result by showing that DAG partitioning is solvable in $2^{O(t^2)} \cdot n$ time if a width-t tree decomposition of the input graph is given. Thus, we improve an algorithm and answer an open question of *S. Alamdari* and *A. Mehrabian* [Lect. Notes Comput. Sci. 7323, 17–28 (2012; Zbl 1342.05109)].

We complement our algorithms by lower bounds on the running time of exact algorithms and on the effectivity of data reduction.

MSC:

05C70	Edge subsets with special properties (factorization, matching, partitioning, covering and packing, etc.)	Cited in **2** Documents
05C85	Graph algorithms (graph-theoretic aspects)	
05C82	Small world graphs, complex networks (graph-theoretic aspects)	
05C12	Distance in graphs	

Keywords:

NP-hard problem; graph algorithms; polynomial-time data reduction; multiway cut; linear-time algorithms; algorithm engineering; evaluating heuristics

Full Text: DOI

References:

[1] Alamdari, S.; Mehrabian, A., On a DAG partitioning problem, (Proceedings of the 9th International Workshop on Algorithms and Models for the Web Graph, (WAW'12), Lecture Notes in Computer Science, vol. 7323, (2012), Springer), 17-28 · Zbl 1342.05109

[2] Barabási, A.-L.; Albert, R., Emergence of scaling in random networks, Science, 286, 5439, 509-512, (1999) · Zbl 1226.05223

[3] van Bevern, R., Towards optimal and expressive kernelization for d-hitting set, Algorithmica, 70, 1, 129-147, (2014) · Zbl 1314.68167

[4] van Bevern, R.; Bredereck, R.; Chopin, M.; Hartung, S.; Hüffner, F.; Nichterlein, A.; Suchý, O., Parameterized complexity of DAG partitioning, (Proceedings of the 8th International Conference on Algorithms and Complexity, (CIAC'13), Lecture Notes in Computer Science, Number 7878, (2013), Springer), 49-60 · Zbl 1382.68127

[5] van Bevern, R.; Hartung, S.; Kammer, F.; Niedermeier, R.; Weller, M., Linear-time computation of a linear problem kernel for dominating set on planar graphs, (Proceedings of the 6th International Symposium on Parameterized and Exact Computation, (IPEC'11), Lecture Notes in Computer Science, vol. 7112, (2012), Springer), 194-206 · Zbl 1352.68119

[6] Bodlaender, H. L., A linear-time algorithm for finding tree-decompositions of small treewidth, SIAM J. Comput., 25, 6, 1305-1317, (1996) · Zbl 0864.68074

[7] Bodlaender, H. L., Kernelization: new upper and lower bound techniques, (Proceedings of the 4th International Workshop on Parameterized and Exact Computation, (IWPEC'09), Lecture Notes in Computer Science, vol. 5917, (2009), Springer), 17-37 · Zbl 1273.68158

[8] Bodlaender, H. L.; Jansen, B. M.P.; Kratsch, S., Kernelization lower bounds by cross-composition, SIAM J. Discrete Math., 28, 1, 277-305, (2014) · Zbl 1295.05222

[9] Cormen, T. H.; Leiserson, C. E.; Rivest, R. L.; Stein, C., Introduction to algorithms, (2001), MIT Press · Zbl 1047.68161

[10] Cygan, M.; Fomin, F. V.; Kowalik, L.; Lokshtanov, D.; Marx, D.; Pilipczuk, M.; Pilipczuk, M.; Saurabh, S., Parameterized algorithms, (2015), Springer · Zbl 1334.90001

[11] Dahlhaus, E.; Johnson, D. S.; Papadimitriou, C. H.; Seymour, P. D.; Yannakakis, M., The complexity of multiterminal cuts, SIAM J. Comput., 23, 4, 864-894, (1994) · Zbl 0809.68075

[12] Downey, R. G.; Fellows, M. R., Fundamentals of parameterized complexity, (2013), Springer · Zbl 1358.68006

[13] Fafianie, S.; Kratsch, S., A shortcut to (Sun)flowers: kernels in logarithmic space or linear time, (Proceedings of the 40th International Symposium on Mathematical Foundations of Computer Science, (MFCS'15), Lecture Notes in Computer Science, vol. 9235, (2015), Springer), 299-310 · Zbl 06482817

[14] Fellows, M. R.; Jansen, B. M.P.; Rosamond, F. A., Towards fully multivariate algorithmics: parameter ecology and the deconstruction of computational complexity, European J. Combin., 34, 3, 541-566, (2013) · Zbl 1448.68465

[15] Flum, J.; Grohe, M., Parameterized complexity theory, (2006), Springer

[16] Fortnow, L.; Santhanam, R., Infeasibility of instance compression and succinct PCPs for NP, J. Comput. System Sci., 77, 1, 91-106, (2011) · Zbl 1233.68144

[17] Guo, J.; Niedermeier, R., Invitation to data reduction and problem kernelization, ACM SIGACT News, 38, 1, 31-45, (2007)

[18] Hagerup, T., Simpler linear-time kernelization for planar dominating set, (Proceedings of the 6th International Symposium on Parameterized and Exact Computation, (IPEC'11), Lecture Notes in Computer Science, volume 7112, (2012), Springer), 181-193 · Zbl 1352.68109

[19] Impagliazzo, R.; Paturi, R., On the complexity of k-SAT, J. Comput. System Sci., 62, 2, 367-375, (2001) · Zbl 0990.68079

[20] Impagliazzo, R.; Paturi, R.; Zane, F., Which problems have strongly exponential complexity?, J. Comput. System Sci., 63, 4, 512-530, (2001) · Zbl 1006.68052

[21] Jeong, H.; Néda, Z.; Barabási, A. L., Measuring preferential attachment in evolving networks, Europhys. Lett., 61, 4, 567-572, (2003)

[22] Kammer, F., A linear-time kernelization for the rooted k-leaf outbranching problem, Discrete Appl. Math., 193, 126-138, (2015) · Zbl 1317.05072

[23] Karger, D. R.; Klein, P.; Stein, C.; Thorup, M.; Young, N. E., Rounding algorithms for a geometric embedding of minimum multiway cut, Math. Oper. Res., 29, 3, 436-461, (2004) · Zbl 1082.90149

[24] Kloks, T., (Treewidth. Computations and Approximations, Lecture Notes in Computer Science, vol. 842, (1994), Springer) · Zbl 0825.68144

[25] Komusiewicz, C.; Niedermeier, R., New races in parameterized algorithmics, (Proceedings of the 37th International Symposium on Mathematical Foundations of Computer Science, (MFCS'12), Lecture Notes in Computer Science, vol. 7464, (2012), Springer), 19-30 · Zbl 1365.68286

[26] Leskovec, J.; Backstrom, L.; Kleinberg, J., Meme-tracking and the dynamics of the news cycle, (Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (KDD'09), (2009), ACM), 497-506

[27] Lohr, S., Study measures the chatter of the news cycle, New York Times, B1, (2009), July 13th, New York edition

[28] Niedermeier, R., Invitation to fixed-parameter algorithms, (2006), Oxford University Press · Zbl 1095.68038

[29] Niedermeier, R., Reflections on multivariate algorithmics and problem parameterization, (Proceedings of the 27th International Symposium on Theoretical Aspects of Computer Science, (STACS'10), LIPIcs, vol. 5, (2010), Schloss Dagstuhl-Leibniz-Zentrum für Informatik), 17-32 · Zbl 1230.68096

[30] Niedermeier, R.; Rossmanith, P., A general method to speed up fixed-parameter-tractable algorithms, Inform. Process. Lett., 73, 3-4, 125-129, (2000) · Zbl 1014.68064

[31] Price, D. D.S., A general theory of bibliometric and other cumulative advantage processes, J. Am. Soc. Inf. Sci., 27, 5, 292-306, (1976)

[32] Protti, F.; Dantas˜da Silva, M.; Szwarcfiter, J., Applying modular decomposition to parameterized cluster editing problems, Theory Comput. Syst., 44, 1, 91-104, (2009) · Zbl 1179.68111

[33] C. Suen, S. Huang, C. Eksombatchai, R. Sosic, J. Leskovec, NIFTY: A system for large scale information flow tracking and clustering. in: Proceedings of the 22nd International Conference on World Wide Web (WWW'13), pp. 1237-1248. IW3C2, 2013. http://dx.doi.org/10.1145/2488388.2488496.

[34] Xiao, M., Simple and improved parameterized algorithms for multiterminal cuts, Theory Comput. Syst., 46, 4, 723-736, (2010) · Zbl 1213.68472