
io-port 05883236**Dondi, Riccardo; Fertin, Guillaume; Vialette, Stéphane****Complexity issues in vertex-colored graph pattern matching.**

J. Discrete Algorithms 9, No. 1, 82-99 (2011).

Summary: Searching for motifs in graphs has become a crucial problem in the analysis of biological networks. In the context of metabolic network analysis, *V. Lacroix, C.G. Fernandes and M.-F. Sagot* [IEEE/ACM Transactions on Computational Biology and Bioinformatics 3, No. 4, 360-368(2006)] introduced the NP-hard general problem of finding occurrences of motifs in vertex-colored graphs, where a motif \mathcal{M} is a multiset of colors and an occurrence of \mathcal{M} in a vertex-colored graph G , called the target graph, is a subset of vertices that induces a connected graph and the multiset of colors induced by this subset is exactly the motif. Pursuing the line of research pioneered by Lacroix et al. and aiming at dealing with approximate solutions, we consider in this paper the above-mentioned problem in two of its natural optimization forms, referred hereafter as the Min-CC and the Maximum Motif problems. The Min-CC problem seeks for an occurrence of a motif \mathcal{M} in a vertex-colored graph G that induces a minimum number of connected components whereas the Maximum Motif problem is concerned with finding a maximum cardinality submotif $\mathcal{M}' \subseteq \mathcal{M}$ that occurs as a connected motif in G . We prove the Min-CC problem to be APX-hard even in the extremal case where the motif is a set and the target graph is a path. We complement this result by giving a polynomial-time algorithm in case the motif is built upon a fixed number of colors and the target graph is a path. Also, extending the results of *M. Fellows, G. Fertin, D. Hermelin and S. Vialette*, in: Proc. 34th International Colloquium on Automata, Languages and Programming (ICALP), Lect. Notes Comput. Sci. 4596, 340–351(2007; Zbl 1171.68497)], we prove the Min-CC problem to be fixed-parameter tractable when parameterized by the size of the motif, and we give a faster algorithm in case the target graph is a tree. Furthermore, we prove the Min-CC problem for trees not to be approximable within ratio $c \log n$ for some constant $c > 0$, where n is the order of the target graph, and to be W[2]-hard when parameterized by the number of connected components in the occurrence of the motif. Finally, we give an exact exponential-time algorithm for the Min-CC problem in case the target graph is a tree. We prove that the Maximum Motif problem is APX-hard even in the case where the target graph is a tree of maximum degree 3, the motif is actually a set and each color occurs at most twice in the tree. Next, we strengthen this result by proving that the problem is not approximable within factor $2^{\log^\delta n}$, for any constant $\delta < 1$, unless $NP \subseteq DTIME(2^{\text{poly} \log n})$. We complement these results by presenting two fixed-parameter algorithms for the problem, where the parameter is the size of the solution. Finally, we give exact exponential-time algorithms for this problem.

Keywords: graph motifs; vertex-colored graphs; algorithmic complexity; parameterized complexity
doi:10.1016/j.jda.2010.09.002