

Cograph editing: Merging modules is equivalent to editing P_4 s*

Adrian Fritz 

*Computational Biology of Infection Research, Helmholtz Centre for Infection Research,
Inhoffenstraße 7, D-38124 Braunschweig*

Marc Hellmuth 

*School of Computing, University of Leeds,
EC Stoner Building, Leeds LS2 9JT, England*

Peter F. Stadler[†] 

*Bioinformatics Group, Department of Computer Science, Universität Leipzig,
Härtelstrasse 16-18, D-04107 Leipzig, Germany*

Nicolas Wieseke 

*Swarm Intelligence and Complex Systems Group, Department of Computer Science,
Leipzig University, Augustusplatz 10, D-04109 Leipzig, Germany*

Received 23 May 2018, accepted 10 September 2019, published online 27 March 2020

Abstract

The modular decomposition of a graph $G = (V, E)$ does not contain prime modules if and only if G is a cograph, that is, if no quadruple of vertices induces a simple connected path P_4 . The cograph editing problem consists in inserting into and deleting from G a set F of edges so that $H = (V, E \triangle F)$ is a cograph and $|F|$ is minimum. This NP-hard combinatorial optimization problem has recently found applications, e.g., in the context of phylogenetics. Efficient heuristics are hence of practical importance. The simple characterization of cographs in terms of their modular decomposition suggests that instead of editing G one could operate directly on the modular decomposition. We show here that editing the

*We thank the anonymous referees for their helpful comments. This work was funded in part by the Deutsche Forschungsgemeinschaft (DFG), Proj. Nr. 432974470 (to PFS).

[†]PFS is also affiliated with the Interdisciplinary Center for Bioinformatics, the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, the Competence Center for Scalable Data Services and Solutions Dresden-Leipzig, the Leipzig Research Center for Civilization Diseases, and the Centre for Biotechnology and Biomedicine at Leipzig University; the Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany; the Institute for Theoretical Chemistry, University of Vienna, Vienna, Austria; the Center of noncoding RNA in Health and Technology (RTH) at the University of Copenhagen; Facultad de Ciencias of the National University of Colombia in Bogotá, Colombia; and the Santa Fe Institute, Santa Fe, NM.

induced P_4 s is equivalent to resolving prime modules by means of a suitable defined merge operation on the submodules. Moreover, we characterize so-called module-preserving edit sets and demonstrate that optimal pairwise sequences of module-preserving edit sets exist for every non-cograph. This eventually leads to an exact algorithm for the cograph editing problem as well as fixed-parameter tractable (FPT) results when cograph editing is parameterized by the so-called modular-width. In addition, we provide two heuristics with time complexity $O(|V|^3)$, resp., $O(|V|^2)$.

Keywords: Cograph editing, modular decomposition, module merge, prime modules, P_4 .

Math. Subj. Class.: 05C75, 05C05, 92B10

Urejanje kografov: združevanje modulov je ekvivalentno urejanju poti P_4 *

Adrian Fritz 

*Computational Biology of Infection Research, Helmholtz Centre for Infection Research,
Inhoffenstraße 7, D-38124 Braunschweig*

Marc Hellmuth 

*School of Computing, University of Leeds,
EC Stoner Building, Leeds LS2 9JT, Anglija*

Peter F. Stadler[†] 

*Bioinformatics Group, Department of Computer Science, Universität Leipzig,
Härtelstrasse 16-18, D-04107 Leipzig, Nemčija*

Nicolas Wieseke 

*Swarm Intelligence and Complex Systems Group, Department of Computer Science,
Leipzig University, Augustusplatz 10, D-04109 Leipzig, Nemčija*

Prejeto 23. maja 2018, sprejeto 10. septembra 2019, objavljeno na spletu 27. marca 2020

Povzetek

Modularna dekompozicija grafa $G = (V, E)$ ne vsebuje praštevilskih modulov natanko tedaj, ko je G kograf, kar pomeni, da nobena četverica točk ne inducira enostavne povezane poti P_4 . *Problem urejanja kografov* zahteva, da v graf G bodisi vstavimo bodisi iz njega odstranimo povezave iz množice F , tako da je $H = (V, E \triangle F)$ kograf, $|F|$ pa minimalen. Ta NP-težak kombinatorično-optimizacijski problem je nedavno našel aplikacije, npr. v filogenetiki. Učinkovita heuristika zanj je torej v praksi zelo pomembna. Enostavna karakterizacija kografov v smislu njihove modularne dekompozicije omogoča, da lahko,

*Zahvaljujemo se anonimnim recenzentom za njihove koristne pripombe. To delo je bilo delno financirano s strani Deutsche Forschungsgemeinschaft (DFG), proj. št. 432974470 (PFS-ju).

[†]PFS je povezan tudi z naslednjimi institucijami: Interdisciplinary Center for Bioinformatics, the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, the Competence Center for Scalable Data Services and Solutions Dresden-Leipzig, the Leipzig Research Center for Civilization Diseases, the Centre for Biotechnology and Biomedicine at Leipzig University; the Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany; the Institute for Theoretical Chemistry, University of Vienna, Vienna, Austria; the Center of noncoding RNA in Health and Technology (RTH) at the University of Copenhagen; Facultad de Ciencias of the National University of Colombia in Bogotá, Colombia; the Santa Fe Institute, Santa Fe, NM.

namesto da bi urejali graf G , delujemo neposredno na njegovo modularno dekompozicijo. V zvezi s tem pokažemo, da je urejanje induciranih poti P_4 ekvivalentno reševanju praštevilskih modulov s pomočjo primerno definirane operacije združevanja podmodulov. Poleg tega karakteriziramo t.i. *urejevalne množice, ki ohranjajo module*, in dokažemo, da za vsak graf, ki ni kograf, obstajajo optimalna parna zaporedja urejevalnih množic, ki ohranjajo module. To naposled vodi k eksaktnemu algoritmu za problem urejanja kografov, pa tudi k sledljivim rezultatom s fiksnim parametrom (FPT) v primeru, ko je urejanje kografa parametrizirano s t.i. *modularno širino*. Poleg tega podamo dve heuristiki, katerih časovni kompleksnosti sta $O(|V|^3)$ oz. $O(|V|^2)$.

Ključne besede: Urejanje kografov, modularna dekompozicija, združevanje modulov, praštevilski moduli, P_4 .

Math. Subj. Class.: 05C75, 05C05, 92B10
