The Journal of Supercomputing (2022) 78:1642–1671
https://doi.org/10.1007/s11227-021-03938-7

Efficient parallel edge-centric approach for relaxed graph pattern matching

Sarra Bouhenni[1,2,3] · Saïd Yahiaoui[2] · Nadia Nouali-Taboudjemat[2] · Hamamache Kheddouci[3]

Accepted: 7 June 2021 / Published online: 15 June 2021

Abstract

Prior algorithms on graph simulation for distributed graphs are not scalable enough as they exhibit heavy message passing. Moreover, they are dependent on the graph partitioning quality that can be a bottleneck due to the natural skew present in real-world data. As a result, their degree of parallelism becomes limited. In this paper, we propose an efficient parallel edge-centric approach for distributed graph pattern matching. We design a novel distributed data structure called *ST* that allows a fine-grain parallelism, and hence guarantees linear scalability. Based on *ST*, we develop a parallel graph simulation algorithm called *PGSim*. Furthermore, we propose *PDSim*, an edge-centric algorithm that efficiently evaluates dual simulation in parallel. *PDSim* combines *ST* and *PGSim* in a Split-and-Combine approach to accelerate the computation stages. We prove the effectiveness and efficiency of these propositions through theoretical guarantees and extensive experiments on massive graphs. The achieved results confirm that our approach outperforms existing algorithms by more than an order of magnitude.

Keywords Graph pattern matching · Subgraph matching · Graph simulation · Dual simulation · Massive graph · Parallel algorithm

✉ Sarra Bouhenni
cs_bouhenni@esi.dz

Saïd Yahiaoui
syahiaoui@cerist.dz

Nadia Nouali-Taboudjemat
nnouali@cerist.dz

Hamamache Kheddouci
hamamache.kheddouci@univ-lyon1.fr

[1] Ecole nationale Supérieure d'Informatique, BP M68, 16309 Oued Smar, Algerie

[2] CERIST, Centre de Recherche sur l'Information Scientifique et Technique, 16030 Ben Aknoun, Algérie

[3] Université de Lyon, CNRS, Université Lyon 1,LIRIS, UMR5205, F-69622 Lyon, France