Karlsruher Institut für Technologie
Fakultät für Informatik

Lehrstuhl Algorithm Engineering
Prof. Dr. Peter Sanders

HiWi-Tätigkeit

KaGen: Karlsruhe Graph Generation

Description

Network generators serve as a tool to alleviate the need for synthethic instances with controllable parameters by algorithm developers and researchers. However, many generators fail to provide instances on a massive scale due to their sequential nature or resource constraints.

KaGen [1] is a collection of novel generators [2, 3, 4] for a variety of network models commonly found in practice. By making use of pseudorandomization and divide-and-conquer schemes, the generators follow a communication-free paradigm. The resulting generators are often embarrassingly parallel and have a near optimal scaling behavior. This allows the generation of instances with up to 2^{43} vertices and 2^{47} edges in less than 22 minutes on 32 768 cores.

The current C++ implementations [1] of these generators are high-performance, but lack many nice-to-have *usability features*. In this HiWi position, we would therefore like to implement such features.

Task

Implement several usability features on top of the already implemented generators. These include:

- Make the overall implementation more stable by implementing unit tests for the generators and fixing broken edge-cases that arise.
- Implement the generation of meaningful edge weights for the currently exclusively unweighted graphs.
- Implement a second layer of parallelization by using OpenMP in addition to the current MPI-based parallelization.

Requirements

- Very good knowledge of C++.
- Interest in efficient algorithms and data structures.
- Knowledge of CMake and Linux or the ability to quickly and independently familiarize yourself with them.

References

[1] Source code: `github.com/KarlsruheGraphGeneration/KaGen`

[2] D. Funke, S. Lamm, P. Sanders, u.a., “Communication-free Massively Distributed Graph Generation”, in: 2018 IEEE International Parallel and Distributed Processing Symposium (IPDPS).

[3] P. Sanders und C. Schulz, “Scalable Generation of Scale-free Graphs”, in: Information Processing Letters, Band 116, S. 489 – 491, 2016.

[4] P. Sanders, L. Hübschle-Schneider, “Linear Work Generation of R-MAT Graphs”, in: Network Science, Band 8, S. 543 – 550, 2020.

Matthias Schimek, Daniel Seemaier, Niklas Uhl
{schimek, daniel.seemaier, uhl}@kit.edu