A collection of benchmark datasets for evaluating graph layout algorithms*

Sara Di Bartolomeo[1,2][0000−0001−9517−3526], Eduardo Puerta[1][0000−0003−4664−4365], Connor Wilson[1][0000−0002−6936−4078], Tarik Cronvrsanin[1][0000−0002−4397−5532], and Cody Dunne[1][0000−0002−1609−9776]

[1] Northeastern University, Boston, USA
[2] Universität Konstanz, Germany
di-bartolomeo@dbvis.inf.uni-konstanz.de, {puerta.e, wilson.conn, t.crnovrsanin, c.dunne}@northeastern.edu

Abstract. We built a website to help graph drawing researchers find benchmark datasets to use for evaluating graph layout algorithms. Find it here: https://visdunneright.github.io/gd_benchmark_sets/. The datasets and supplemental materials are also available at https://osf.io/j7ucv/.

Benchmarking is a crucial aspect of computer science, as it allows researchers, developers, and engineers to compare the performance of various systems, algorithms, or hardware. A benchmark is a standardized test or set of tests used to measure and compare the performance of hardware, software, or systems under specific conditions. Benchmarking aims to provide objective and consistent metrics that allow for fair comparisons and informed decision-making. Benchmarks are widely used in various fields, including computer hardware evaluation, software optimization, and system performance analysis. In all these fields, benchmarking provides a standardized and objective way to compare and assess the performance of different systems, algorithms, or software implementations. It aids in making informed decisions about which solution best suits a specific use case or requirement.

The same is true for the field of graph drawing, and in particular, for studying the performance and results of graph layout algorithms. Benchmark datasets can provide a standardized set of graphs with known properties and characteristics. These graphs can vary in size, density, connectivity, and structure. Researchers can objectively compare their performance or the quality of their results by applying various graph layout algorithms to the same benchmark dataset.

Because of our own challenges in finding appropriate benchmark sets to evaluate layout algorithms that we developed, we built a collection of benchmark datasets used in previous graph layout algorithm papers and a website to peruse the collection we put together.

The key objective of the work we are doing with benchmark datasets is not only aimed at improving the discoverability of these datasets and easing the

* Supported in part by the U.S. National Science Foundation (NSF) under award number IIS-2145382.

running of benchmarks for graph layout algorithms of a vast amount of types and categories, but we also want to place a strong emphasis on the replicability of the experiments that are run. Indeed, reliable access to datasets is fundamental for replicability.

The collection: the information we collected is a by-product of a larger systematic review we conducted related to graph layout algorithms, which included 206 papers—the core of them being the last 7 years of Graph Drawing proceedings, filtering out the papers with no computational evaluations. Our research expanded to include papers from TVCG and CGF, sourced from the IEEE, ACM, and Wiley digital libraries. For each paper, we noted the algorithm features and datasets used. To locate datasets, we checked supplemental materials, searched online, or contacted authors. We sought permissions for dataset redistribution and stored unclaimed or approved datasets for preservation. We respect ownership rights and will remove any dataset upon the owner's request.

The website is accessible at https://visdunneright.github.io/gd_benchmark_sets/. Every dataset is accompanied by:

- Labels describing what graph features can be found in the dataset. Additionally, we offer a summary analysis of the contents of the dataset, including information about the distribution of node degrees, or how many graphs are contained in a given dataset, or how many nodes and edges they have.
- A link to where to find the dataset and what paper was associated with its initial publication (if any).
- A list of papers that have used the dataset that exemplify its use in previous research. Moreover, we include representative images of how the dataset has been used in previous research to provide an immediate impression of how the dataset would look. Additionally, we collected the text descriptions of the dataset in these previous papers, which reports useful information in addition to, in some instances, additional insights obtained by the authors.
- A link to the storage location on OSF, which includes converting the original data format to four common formats: GEXF, GraphML, GML, and JSON.

Conclusion: Benchmarking is an important tool in computer science, especially in graph drawing, where consistent datasets are essential for evaluating algorithms. Addressing the challenge of sourcing these datasets, we have curated a collection from prior studies and created a user-friendly website for accessibility. This endeavor not only streamlines the benchmarking process but also emphasizes the replicability of experiments. The website offers a concise overview of each dataset, its features, and associated research. We hope to assist researchers in efficiently finding the right datasets for their work.