Supplemental Material for
Algorithmic Improvements on Hilbert and Moore Treemaps
for Visualization of Large Tree-structured Datasets

W. Scheibel[1] , C. Weyand[1], J. Bethge[1] , and J. Döllner[1]

[1]Hasso Plattner Institute, Digital Engineering Faculty, University of Potsdam, Germany

Prototype

The prototype is published as an Open Source repository on Github at https://github.com/varg-dev/hilbert-moore-treemap-layouts-prototype. The prototype can be used to reproduce the layouts, the metrics assessment and the performance evaluation. The used datasets are integrated as well. Please refer to the README.md of the repository for more information on how to setup, compile, and use the prototype.

Processing Specifics

The AAR was measured for each dataset and each snapshot separately and averaged over all datasets and snapshots. The four metrics ADC, RPC, AAD, and RDC are measured for all datasets with more than one snapshot. These metrics are computed for each pair of layouts from successive snaphots within a dataset and are averaged over all datasets and pairs. The LD metric is measured for all datasets with more than one snapshot, too. It computes one value for a whole dataset by taking all layouts into account. This value is averaged over all datasets.

Used Datasets

Following is the list of used datasets and basic statistics. If a dataset spans multiple snapshots, the number of nodes and the max depth are derived from the snapshot with the maximum number of leaf nodes. The used weight metric is either *Lines of Code* (LoC) or *Number of Stars* (NoS).

Name	Github Identifier	Weight	# Nodes	# Leaf Nodes	Max Depth	# Snapshots
Calculator	microsoft/calculator	LoC	371	334	6	510
CPP Rest SDK	microsoft/cpprestsdk	LoC	579	446	9	930
glbinding	cginternals/glbinding	LoC	740	639	9	644
glfw	glfw/glfw	LoC	117	105	4	3 282
globjects	cginternals/globjects	LoC	454	407	9	891
GSL	microsoft/GSL	LoC	21	18	3	354
Open Source Portal	microsoft/opensource-portal	LoC	413	334	6	512
STL	microsoft/STL	LoC	1 368	800	7	487
TensorFlow.js	tensorflow/tfjs	LoC	2 060	1 894	12	1 216
The Algorithms C++	TheAlgorithms/C-Plus-Plus	LoC	239	218	4	529
VFS For Git	microsoft/VFSForGit	LoC	943	790	7	609
webgl-operate	cginternals/webgl-operate	LoC	202	183	4	550
Kubernetes	kubernetes/kubernetes	LoC	16 963	13 175	16	1
Firefox	mozilla/gecko-dev	LoC	114 364	105 579	16	1
GitHub Project Landscape	–	NoS	2 391 549	1 851 123	3	1

Example Layouts

Layout examples for the mid-sized and large-sized software repositories expect the Github Landscape dataset. The used subdivision algorithm is the min-variance weight partitioning with the Hilbert space-filling curve. The color mapping is the node index in the dataset – not the position in the tree – to a grayscale color gradient. Note: the layout for the Github landscape is currently not processable by PDF viewers and printers.

Calculator

CPP Rest SDK

glbinding

glfw

globjects

GSL

Open Source Portal

STL

Tensorflow.js

The Algorithms C++

VFS for Git

webgl-operate

Kubernetes

Firefox