

KONECT Cloud

Large Scale Network Mining in the Cloud

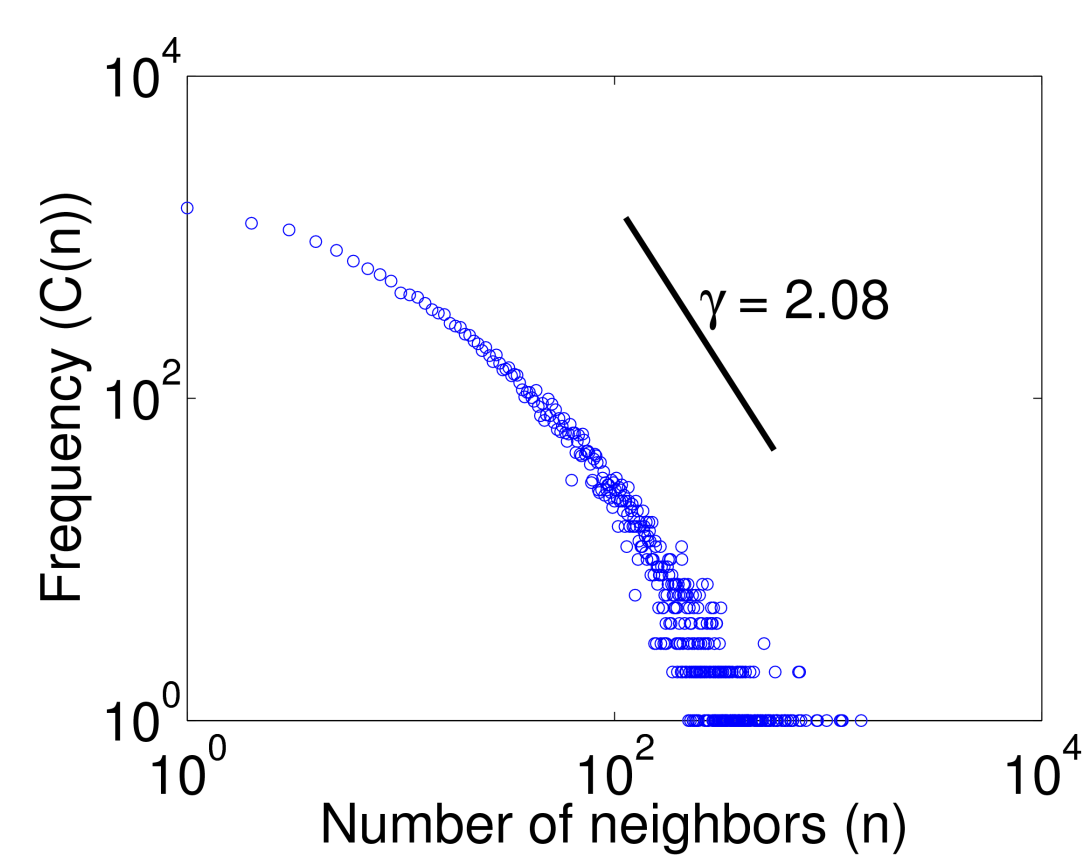
KONECT – Koblenz Network Collection

Framework of network datasets and code for analysis, targeting large-scale networks from the Web.

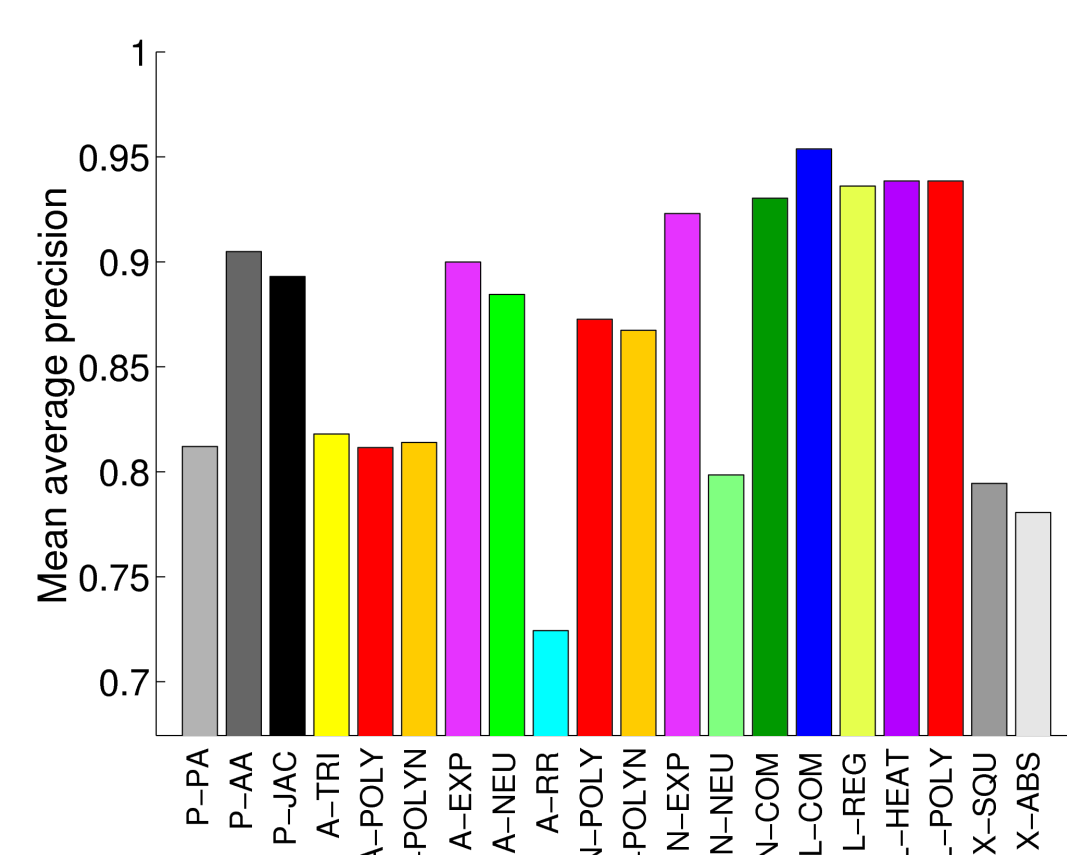
konect.uni-koblenz.de

All code available under GPLv3 / cc-by-sa

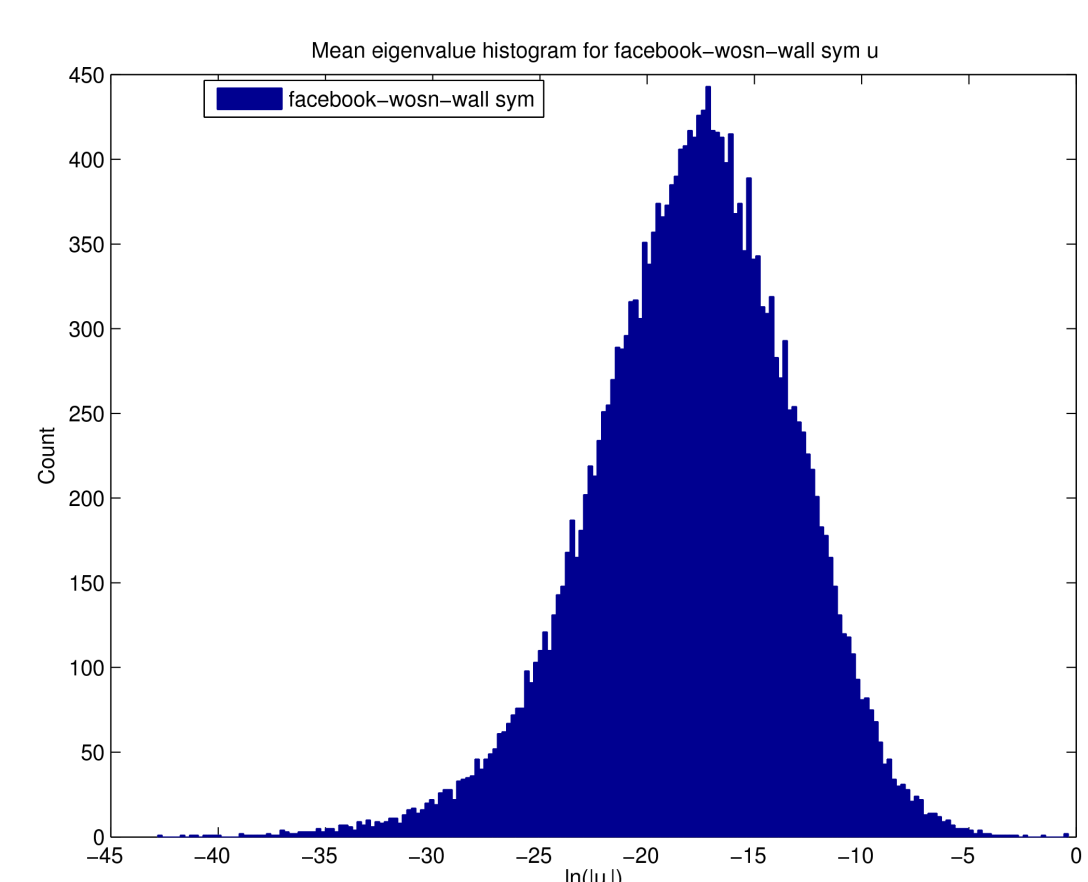
Visual Network Analysis



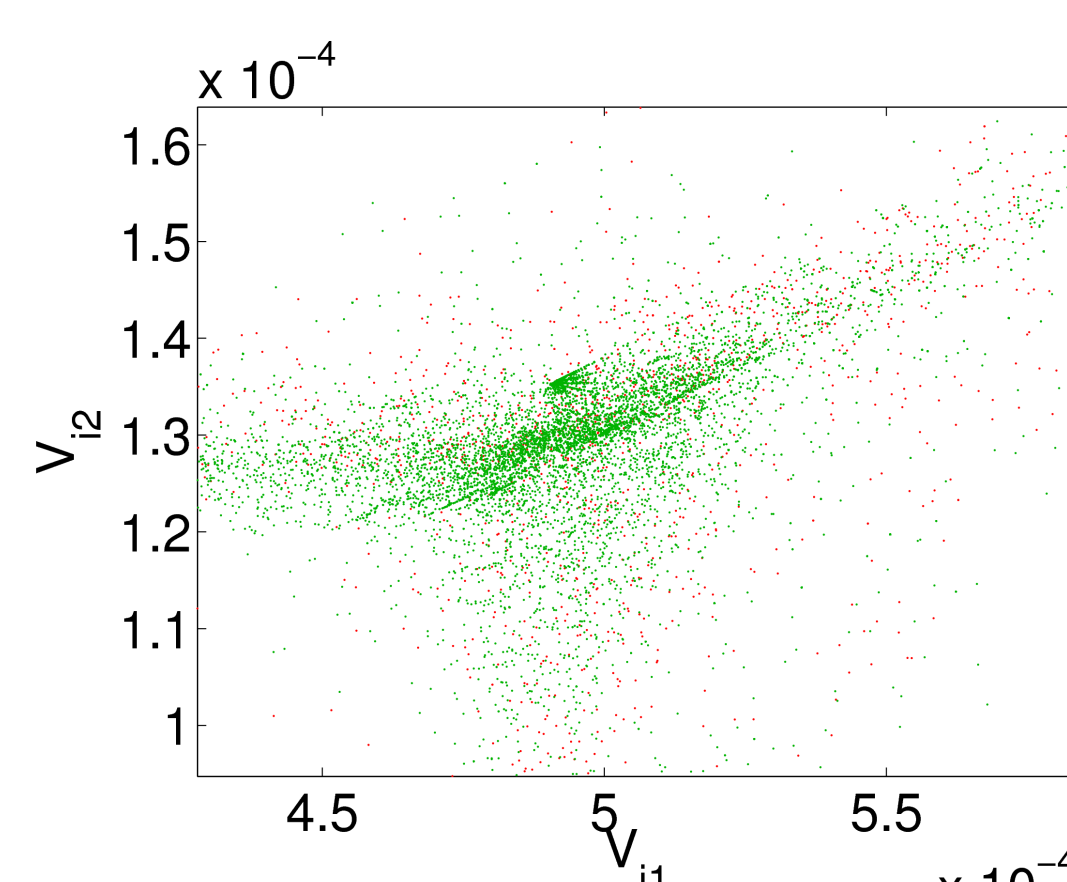
Degree distribution



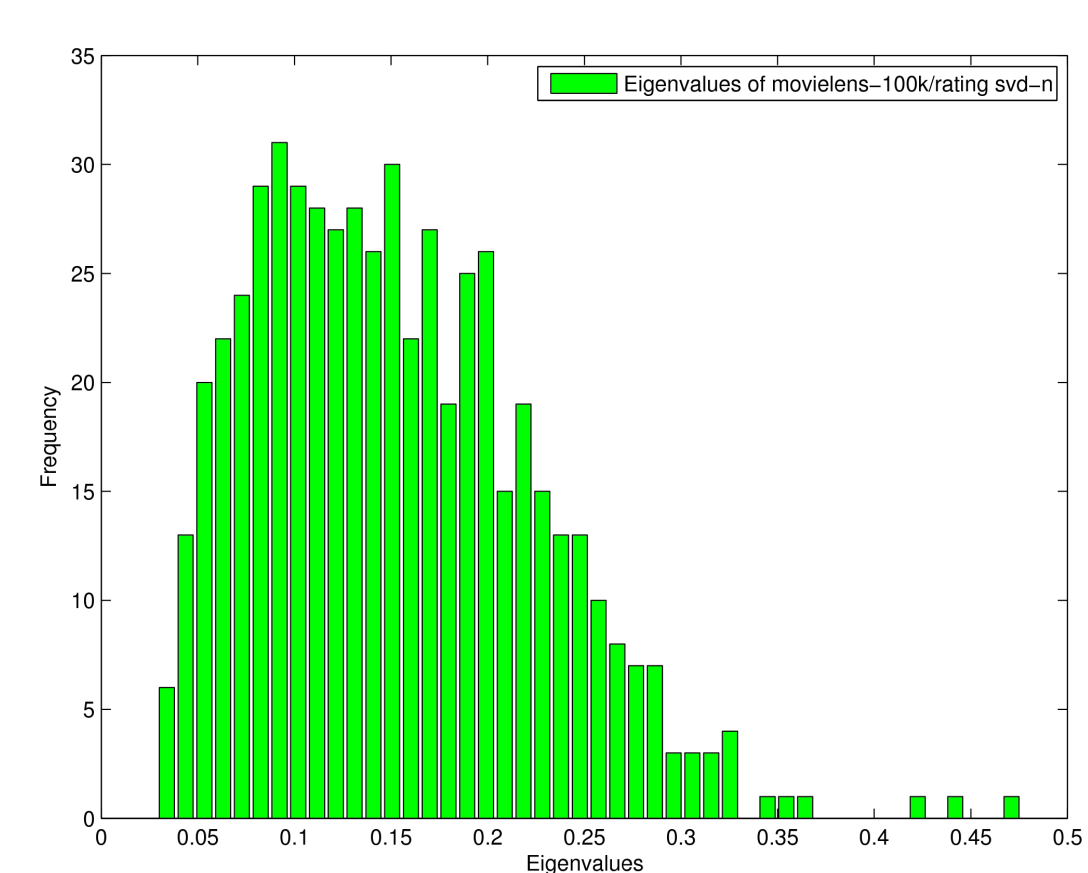
Link prediction results



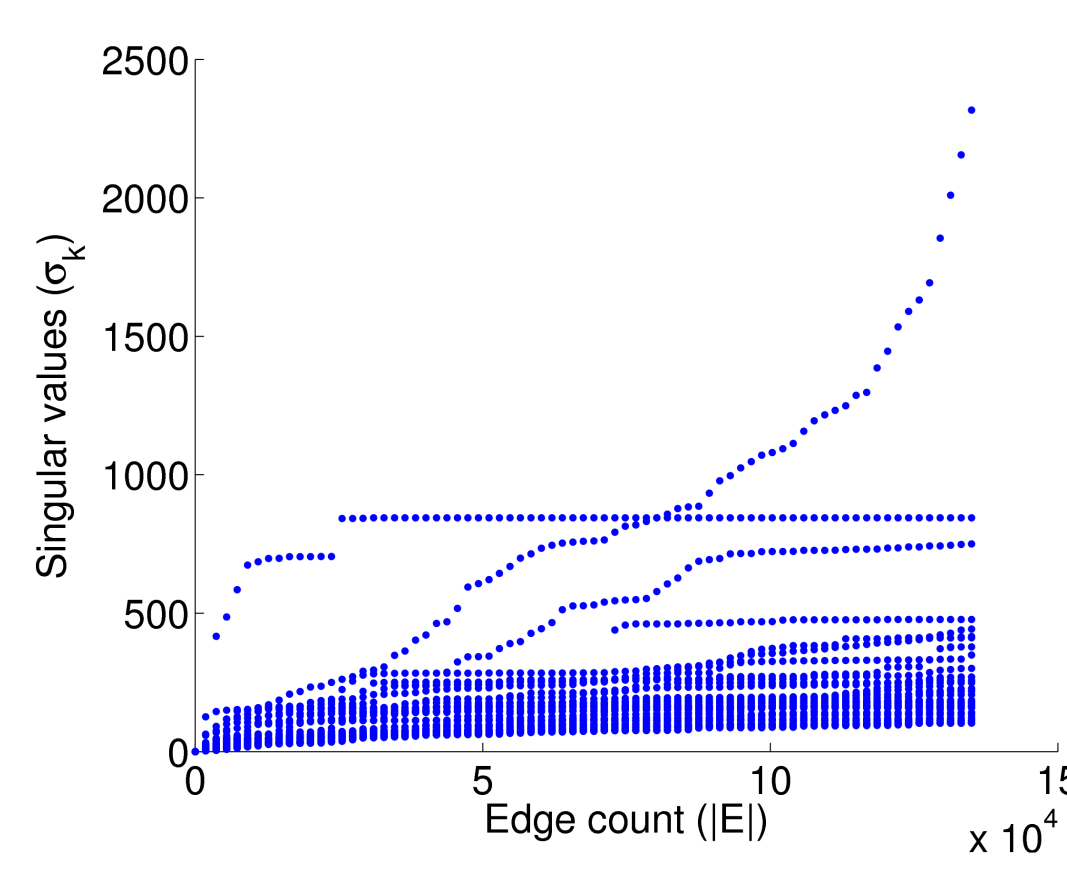
Eigenvector centrality



Scatter plots of nodes



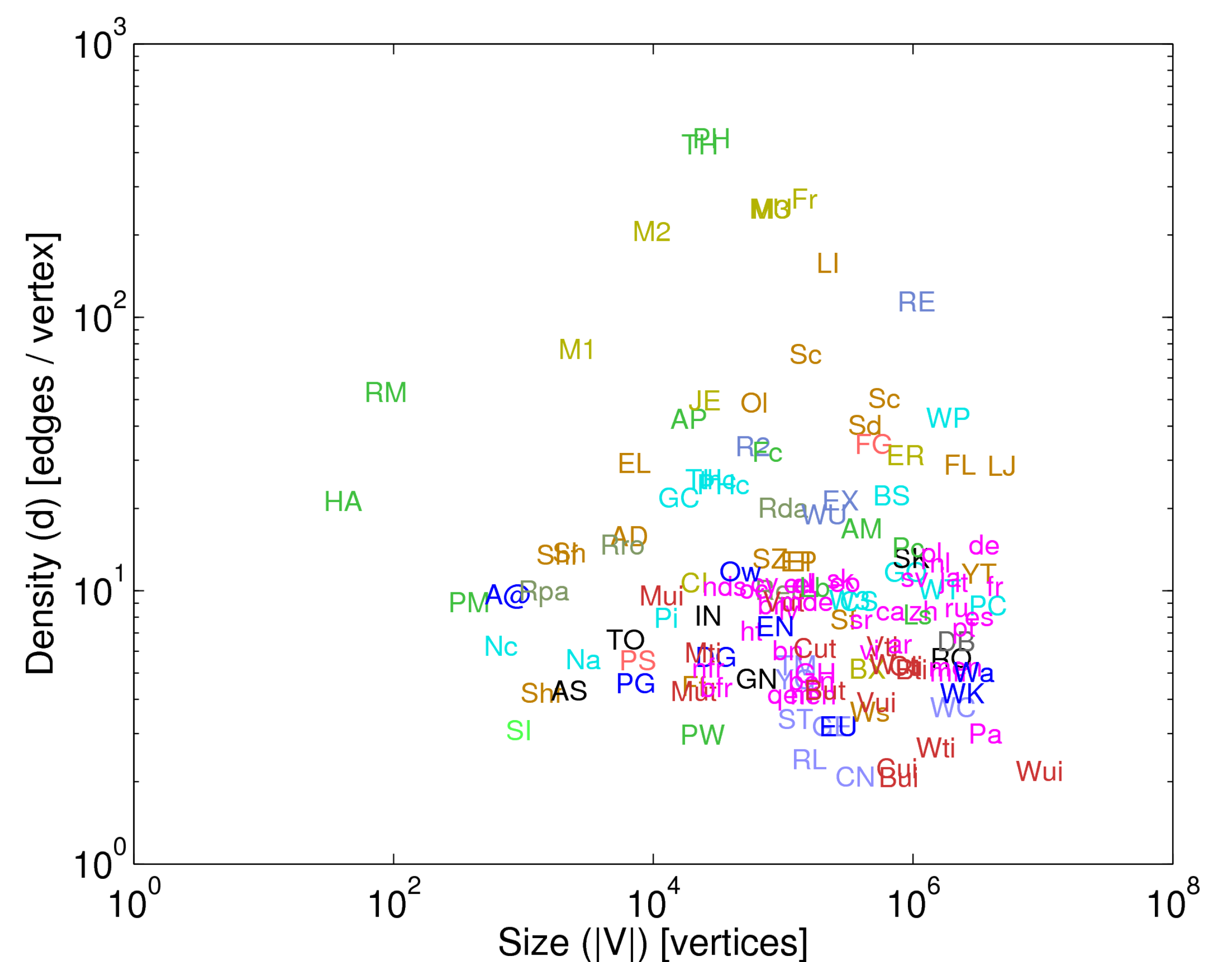
Spectral distribution



Eigenvectors over time

147 Network Datasets

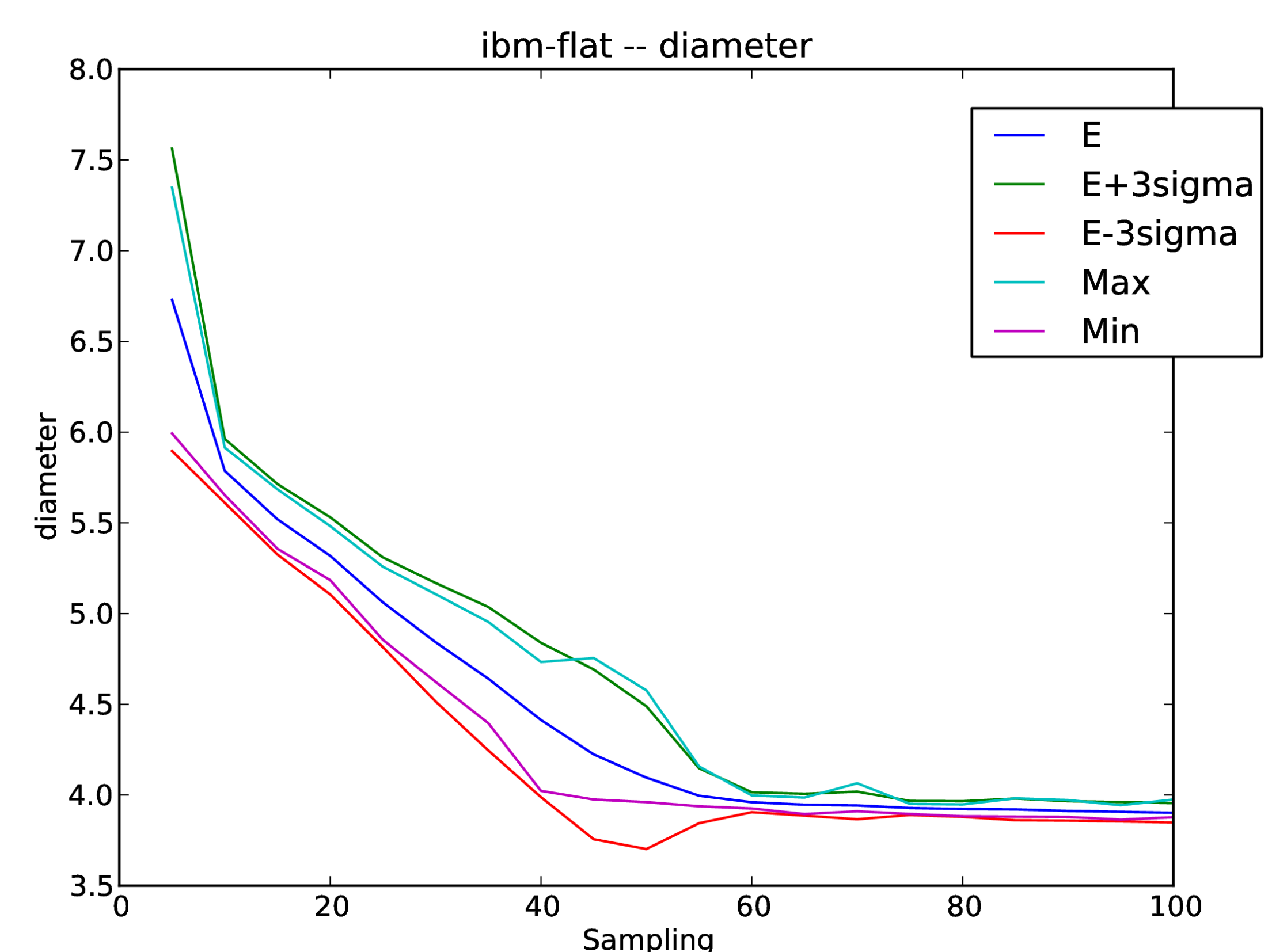
authorship communication co-occurrence
features folksonomy interaction physical
ratings reference semantic social trust



We provide: Code for download & conversion into Octave/Matlab format

Results at Future SOC Lab: Graph Sampling

How are network statistics affected by *downsampling*, i.e. keeping on $N\%$ of all edges?



Example: The effective network diameter does *not* change by a sampling of $\geq 60\%$.