A scale-free benchmark graphs
for overlapping community detection algorithms

Jean-Gabriel Young§, Laurent Hébert-Dufresne†, Edward Laurence§ & Louis J. Dubé§.

§Département de physique, de génie physique et d’optique, Université Laval, Québec, QC, Canada.
†Santa Fe Institute, Santa Fe, NM 87501, USA

Summary

We introduce a large class of scale-free benchmark graphs for overlapping community detection algorithms.

This organic approach to benchmarking allows us
• to generate a wide range of community structures;
• to identify qualitative structural regimes easily;
• to analyze the strengths and weaknesses of an algorithm at a glance.

Graph properties

The structural properties of the graph (e.g. clustering coefficient, degree) are functions of the input parameters 
(p, q, r, N), rather than imposed directly. These properties vary smoothly with the parameters.

Network structure   Community structure   Running time [sec.]

Average degree     Largest community    r = 1, N = 5000
Clustering coefficient  Highest memberships   r = 10, N = 5000
Degree assortativity     Excess density         r = 1, N = 25 000
Maximal coreness      Average overlap        r = 10, N = 25 000

A subset of properties for fixed values of (N, r).
Similar behaviors are obtained for all (N, r) pairs.

Relationship between (p, q, N, r) and the expected running time.

Case study: OSLOM

We applied the OSLOM algorithm [§] to our benchmark for multiple (p, q) pairs (fixed N, r).

OSLOM performs poorly whenever p, q are small; i.e. for dense, clustered networks with large communities (left).

More importantly, we observe transitions in detectability along multiple trajectories in the configuration space (right).

Accuracy (NMI)

OSLOM

Accuracy on trajectories

OSLOM

Conclusion and future work

We have introduced a large class of scale-free benchmark graphs for overlapping community detection algorithms. This organic approach to benchmarking allows us to generate a wide range of community structures; to identify qualitative structural regimes easily; and to analyze the strengths and weaknesses of an algorithm at a glance. We applied the OSLOM algorithm to our benchmark for multiple (p, q) pairs (fixed N, r). OSLOM performs poorly whenever p, q are small; i.e. for dense, clustered networks with large communities (left). More importantly, we observe transitions in detectability along multiple trajectories in the configuration space (right).