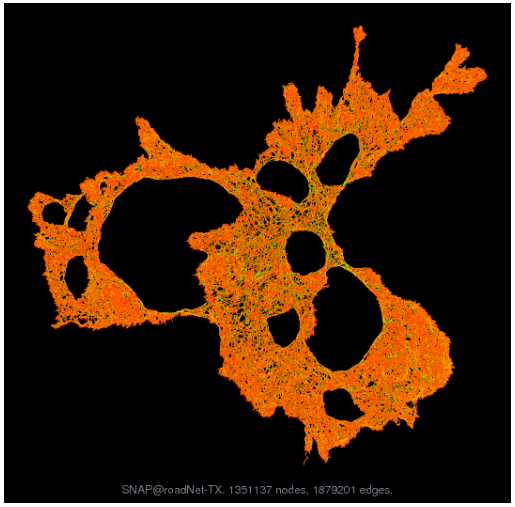
	Facebook[1]	TX roads[2]
Vertices	63731	1393383
Edges (both undir.)	817035	1921660
Triangles	3500542	82869
Triangles per node	164.780	0.178
Mean degree	25.640	2.758
Density	0.0004023	0.0000002
Conn. components	144	13890
Mean component size	442.576	100.315
Global clustering coeff.	0.1477	0.0602



Question

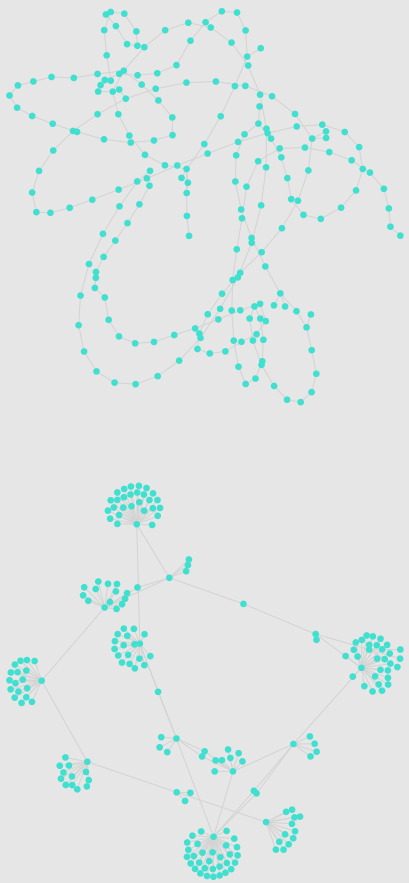
A road network is topologically fundamentally different from a social one. How do different sampling strategies react to these differences?

[1] http://konect.uni-koblenz.de/networks/facebook-wosn-links
[2] http://konect.uni-koblenz.de/networks/roadNet-TX

Comparison of *random walk* and *forest fire* (see here[3]), by analysing some graph measures on samples.

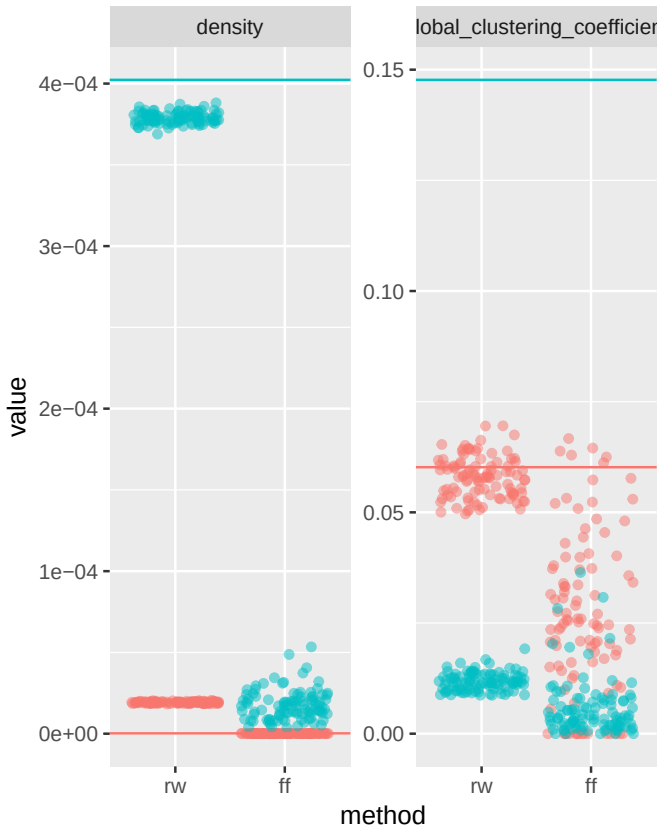
Texas

Facebook

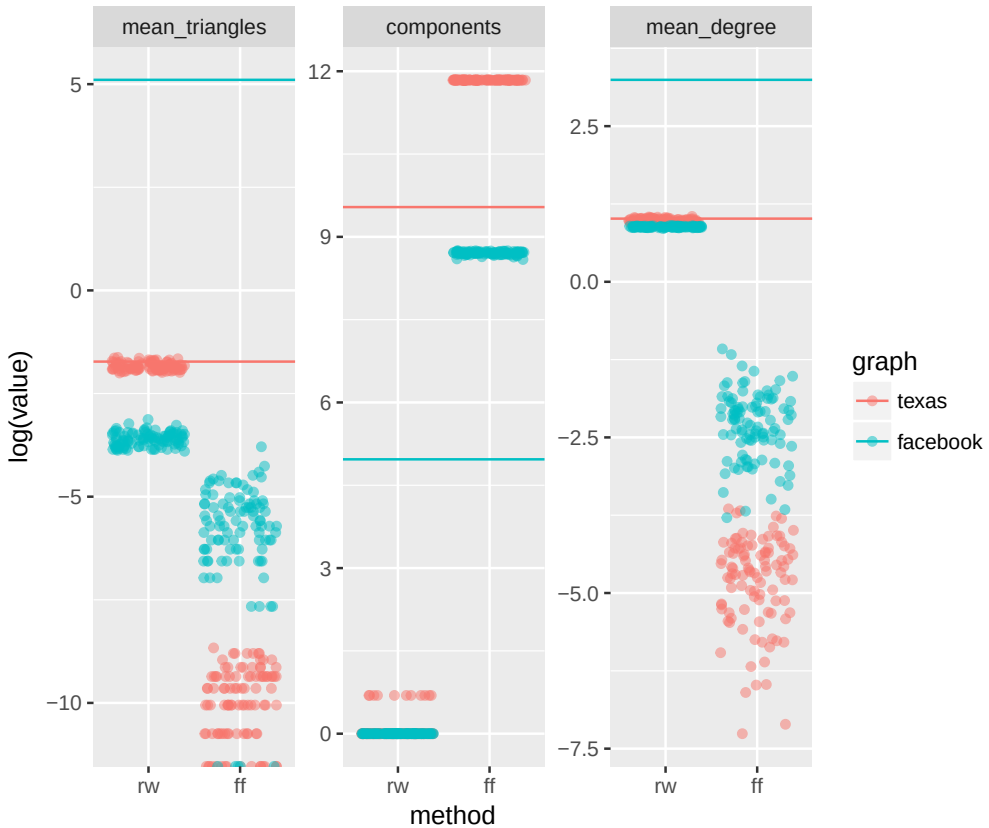


[3] Leskovec, J; Faloutsos, C. "Sampling from Large Graphs." In: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 631–636. ACM, 2006.

Some meaningful measures by sampling 100 times from both data sets, using 10 % of the network[4].



Magnitude mostly preserved



Magnitudes far off, but order still mostly preserved

[4] Code, data, and statistics: `https://github.com/phipsgabler/netsci-01`