

Local network community detection with continuous optimization of conductance and weighted kernel k-means – Supplementary material

Twan van Laarhoven
Elena Marchiori
Institute for Computing and Information Sciences
Radboud University, Nijmegen, The Netherlands
tvanlaarhoven@cs.ru.nl, elenam@cs.ru.nl

June 1, 2016

Additional experimental results

The tables and figures below include additional measurements and results on more datasets.

- Table 1: F_1 scores.
- Table 2: Mean size of the recovered community.
- Table 3: Standard deviation of the size of the recovered community.
- Table 4: Mean conductance of the recovered community.
- Table 5: Standard deviation of the conductance of the recovered community.
- Table 6: Average runtime for finding a community.
- Figure 1: F_1 score on SNAP datasets as a function of σ .
- Figure 2: Mean size of the recovered community as a function of σ .
- Figure 3: Mean conductance of the recovered community as a function of σ .

Additionally we have repeated most experiments with sets of 3 seeds rather than 1 seed,

- Table 7: F_1 scores.
- Table 8: Mean size of the recovered community.
- Table 9: Mean conductance of the recovered community.

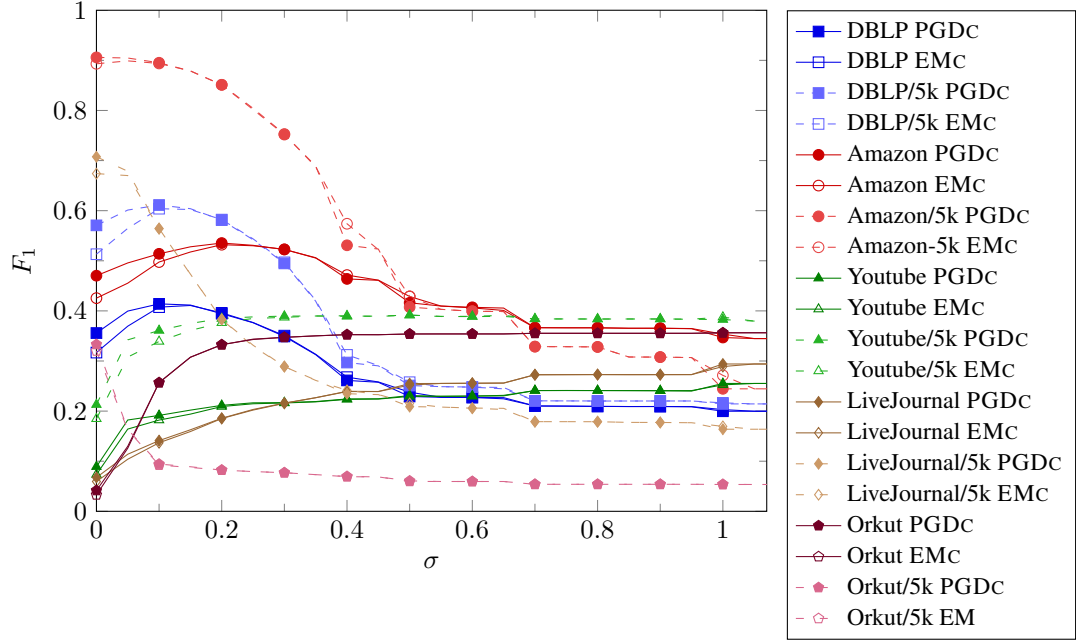


Figure 1: F_1 score as a function of the σ parameter on the SNAP datasets with 1 seed.

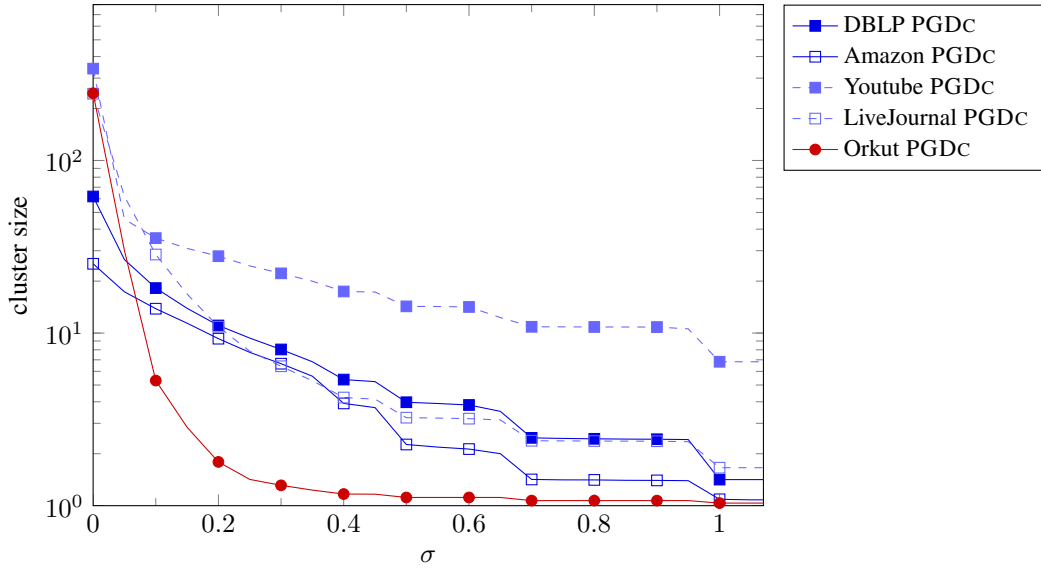


Figure 2: Cluster size as a function of the σ parameter on the SNAP datasets with 1 seed.

Table 1: F_1 scores with 1 seed.

Dataset	Communities	PGDC-0	PGDC-d	EMC-0	EMC-d	YL	HK	PPR
LFR (om=1)		0.967	0.185	0.868	0.187	0.203	0.040	0.041
LFR (om=2)		0.483	0.095	0.293	0.092	0.122	0.039	0.041
LFR (om=3)		0.275	0.085	0.158	0.083	0.110	0.037	0.039
LFR (om=4)		0.178	0.074	0.100	0.072	0.092	0.032	0.034
Karate		0.831	0.472	0.816	0.467	0.600	0.811	0.914
Football		0.792	0.816	0.766	0.805	0.816	0.471	0.283
Pol.Blogs		0.646	0.141	0.661	0.149	0.017	0.661	0.535
Pol.Books		0.596	0.187	0.622	0.197	0.225	0.641	0.663
Flickr		0.098	0.027	0.097	0.027	0.013	0.054	0.118
Amazon	All	0.470	0.522	0.425	0.522	0.493	0.245	0.130
	All, $ C \geq 10$	0.478	0.312	0.490	0.311	0.336	0.428	0.274
	Top 5000	0.906	0.770	0.893	0.771	0.799	0.726	0.539
	Top 5000, $ C \geq 10$	0.810	0.527	0.817	0.526	0.577	0.841	0.753
DBLP	All	0.356	0.369	0.317	0.371	0.341	0.214	0.210
	All, $ C \geq 10$	0.259	0.217	0.238	0.221	0.187	0.176	0.181
	Top 5000	0.571	0.531	0.513	0.531	0.498	0.360	0.348
	Top 5000, $ C \geq 10$	0.491	0.378	0.466	0.378	0.342	0.380	0.359
Youtube	All	0.089	0.251	0.073	0.248	0.228	0.037	0.071
	All, $ C \geq 10$	0.072	0.113	0.070	0.113	0.101	0.015	0.043
	Top 5000	0.213	0.410	0.185	0.408	0.381	0.088	0.156
	Top 5000, $ C \geq 10$	0.091	0.106	0.084	0.108	0.094	0.020	0.048
LiveJournal	All	0.067	0.262	0.059	0.259	0.183	0.035	0.049
	All, $ C \geq 10$	0.096	0.122	0.091	0.121	0.098	0.056	0.062
	Top 5000	0.708	0.663	0.674	0.663	0.557	0.583	0.492
	Top 5000, $ C \geq 10$	0.677	0.585	0.664	0.583	0.434	0.563	0.447
Orkut	All	0.042	0.231	0.033	0.231	0.171	0.057	0.033
	All, $ C \geq 10$	0.067	0.090	0.059	0.091	0.087	0.069	0.045
	Top 5000	0.333	0.150	0.320	0.149	0.115	0.319	0.265
	Top 5000, $ C \geq 10$	0.350	0.077	0.343	0.077	0.048	0.303	0.243
CYC/Gavin 2006		0.474	0.543	0.455	0.543	0.526	0.336	0.294
CYC/Krogan 2006		0.410	0.513	0.364	0.511	0.504	0.229	0.169
CYC/Collins 2007		0.346	0.429	0.345	0.429	0.416	0.345	0.345
CYC/Costanzo 2010		0.174	0.355	0.172	0.351	0.314	0.170	0.170
CYC/Hoppins 2011		0.368	0.405	0.368	0.405	0.424	0.368	0.368
CYC/all		0.044	0.459	0.017	0.459	0.425	0.016	0.002

Table 2: Mean community sizes with 1 seed.

Dataset	Communities	PGDC-0	PGDC-d	EMC-0	EMC-d	YL	HK	PPR
LFR (om=1)		52.3	6.2	71.8	6.3	5.9	2410.0	2366.2
LFR (om=2)		93.2	5.6	292.8	6.4	4.9	2404.4	2311.6
LFR (om=3)		104.7	5.6	451.5	6.4	5.0	2399.4	2283.7
LFR (om=4)		108.9	4.9	530.2	5.7	4.9	2389.1	2262.0
Karate		20.0	8.0	24.1	8.0	8.8	16.7	17.1
Football		14.7	9.5	16.2	9.4	8.8	40.5	56.5
Pol.Blogs		515.6	110.1	538.9	118.1	7.2	492.7	1051.1
Pol.Books		37.8	5.2	43.1	5.7	6.7	49.3	53.4
Flickr		639.9	73.0	644.6	73.5	12.9	174.2	1158.1
Amazon	All	25.2	5.6	45.6	5.7	6.4	88.8	20819.9
	All, $ C \geq 10$	26.9	7.7	41.8	8.2	7.2	99.4	17153.9
	Top 5000	10.9	6.5	12.8	6.5	7.1	28.3	4828.6
	Top 5000, $ C \geq 10$	17.6	8.2	18.6	8.3	9.4	30.4	916.9
DBLP	All	61.9	5.4	83.5	6.1	6.0	55.0	24495.0
	All, $ C \geq 10$	79.6	5.7	105.8	6.6	6.2	61.6	28948.1
	Top 5000	26.4	5.4	38.1	5.6	6.2	46.8	9506.8
	Top 5000, $ C \geq 10$	32.9	6.4	45.1	6.7	6.3	53.2	9390.1
Youtube	All	340.6	19.4	474.2	21.3	9.3	147.9	20955.5
	All, $ C \geq 10$	371.2	34.4	486.5	36.8	9.5	296.3	24846.3
	Top 5000	225.8	16.0	314.2	18.3	6.9	113.9	10867.7
	Top 5000, $ C \geq 10$	350.0	37.8	472.6	42.3	10.0	150.2	21833.7
LiveJournal	All	243.7	5.5	309.3	5.9	10.8	153.2	3428.7
	All, $ C \geq 10$	288.2	8.0	352.9	8.7	11.2	162.1	3978.8
	Top 5000	64.8	26.5	76.3	26.8	13.0	100.0	549.8
	Top 5000, $ C \geq 10$	86.8	36.9	97.9	36.7	15.1	122.8	697.2
Orkut	All	245.1	17.9	344.7	19.1	11.1	212.0	1634.0
	All, $ C \geq 10$	257.2	18.6	331.3	20.4	10.9	220.8	1749.8
	Top 5000	256.9	23.4	284.0	24.1	9.1	264.8	1607.4
	Top 5000, $ C \geq 10$	260.5	22.4	284.5	22.6	9.6	279.1	1594.3
CYC/Gavin 2006		19.7	3.4	34.3	3.5	3.1	236.7	621.9
CYC/Krogan 2006		48.4	7.0	138.8	9.4	3.7	723.8	1756.3
CYC/Collins 2007		202.9	19.5	207.2	19.5	2.6	192.0	189.4
CYC/Costanzo 2010		540.2	58.1	564.2	66.5	5.9	1058.8	942.9
CYC/Hoppins 2011		229.9	110.1	235.5	110.4	4.3	295.2	295.2
CYC/all		657.5	16.0	841.9	17.0	9.6	2795.2	5786.0

Table 3: Standard deviation of community sizes with 1 seed.

Dataset	Communities	PGDC-0	PGDC-d	EMC-0	EMC-d	YL	HK	PPR
LFR (om=1)		29.1	9.6	49.3	9.9	5.0	41.2	36.9
LFR (om=2)		60.1	13.8	149.0	19.4	3.4	61.6	28.7
LFR (om=3)		66.7	14.4	182.0	19.3	3.4	75.4	29.6
LFR (om=4)		74.0	13.4	191.7	20.0	3.5	87.4	28.3
Karate		7.7	7.6	7.1	7.7	4.4	2.5	2.1
Football		9.9	3.1	12.4	3.1	2.4	19.4	9.9
Pol.Blogs		276.0	238.1	273.4	249.1	11.2	222.2	424.0
Pol.Books		19.4	6.4	17.7	7.9	4.9	10.6	3.2
Flickr		255.3	218.4	254.4	219.0	8.4	161.5	648.2
Amazon	All	42.0	5.7	81.2	6.9	4.1	207.8	26078.2
	All, $ C \geq 10$	28.8	9.1	63.9	11.3	5.1	263.4	25419.3
	Top 5000	9.3	5.1	19.9	5.2	4.4	84.7	13773.8
	Top 5000, $ C \geq 10$	10.7	6.7	12.4	6.8	6.2	21.5	6087.3
DBLP	All	97.5	9.4	109.2	17.2	3.7	118.0	37843.3
	All, $ C \geq 10$	108.6	11.3	120.9	18.8	4.1	147.4	40170.2
	Top 5000	41.1	4.5	53.8	7.1	3.1	80.2	24683.8
	Top 5000, $ C \geq 10$	41.8	6.1	58.0	9.3	3.8	99.4	24779.2
Youtube	All	294.5	75.6	361.6	84.3	11.6	864.4	24723.1
	All, $ C \geq 10$	287.0	119.8	348.7	132.9	14.7	1338.2	26044.7
	Top 5000	288.7	73.1	359.8	89.4	11.5	763.9	20851.3
	Top 5000, $ C \geq 10$	287.6	130.0	345.7	143.5	11.3	682.1	25680.0
LiveJournal	All	201.2	23.5	253.1	27.9	8.0	370.2	4768.2
	All, $ C \geq 10$	229.7	35.7	266.7	44.7	8.4	382.9	5127.4
	Top 5000	110.0	52.0	132.8	57.3	15.3	199.3	1503.3
	Top 5000, $ C \geq 10$	126.9	70.2	151.0	70.1	14.8	213.8	1738.9
Orkut	All	214.9	65.1	294.5	76.2	10.1	223.8	1687.5
	All, $ C \geq 10$	199.2	68.4	266.2	88.1	7.1	222.8	1643.8
	Top 5000	186.8	69.3	208.4	70.6	4.8	215.7	2166.2
	Top 5000, $ C \geq 10$	178.4	69.1	195.4	70.6	5.3	209.8	1937.8
CYC/Gavin 2006		51.5	6.0	97.5	6.5	3.1	406.9	676.8
CYC/Krogan 2006		124.9	33.2	254.5	50.9	2.9	974.5	1163.8
CYC/Collins 2007		323.2	111.1	325.4	111.1	3.5	315.9	307.4
CYC/Costanzo 2010		444.3	221.7	442.4	237.5	7.7	857.6	732.9
CYC/Hoppins 2011		418.0	309.9	421.7	309.9	14.9	529.0	529.0
CYC/all		349.7	99.8	240.0	106.1	14.9	892.4	827.8

Table 4: Mean conductance of the recovered communities with 1 seed.

Dataset	Communities	PGDC-0	PGDC-d	EMC-0	EMC-d	YL	HK	PPR
LFR (om=1)		0.301	0.750	0.304	0.749	0.755	0.250	0.273
LFR (om=2)		0.532	0.786	0.541	0.787	0.780	0.315	0.338
LFR (om=3)		0.589	0.793	0.587	0.793	0.781	0.333	0.354
LFR (om=4)		0.604	0.791	0.595	0.792	0.775	0.341	0.359
Karate		0.129	0.460	0.081	0.475	0.327	0.222	0.136
Football		0.277	0.356	0.274	0.362	0.385	0.244	0.155
Pol.Blogs		0.228	0.743	0.212	0.737	0.867	0.229	0.137
Pol.Books		0.140	0.622	0.107	0.611	0.571	0.127	0.065
Flickr		0.777	0.937	0.777	0.937	0.951	0.864	0.762
Amazon	All	0.181	0.464	0.180	0.463	0.402	0.081	0.053
	All, $ C \geq 10$	0.148	0.465	0.152	0.461	0.417	0.068	0.048
	Top 5000	0.071	0.217	0.076	0.217	0.189	0.039	0.030
	Top 5000, $ C \geq 10$	0.062	0.344	0.060	0.345	0.308	0.021	0.018
DBLP	All	0.246	0.571	0.257	0.565	0.498	0.133	0.147
	All, $ C \geq 10$	0.262	0.626	0.276	0.619	0.543	0.139	0.154
	Top 5000	0.197	0.452	0.211	0.450	0.407	0.110	0.119
	Top 5000, $ C \geq 10$	0.203	0.493	0.207	0.490	0.458	0.107	0.117
Youtube	All	0.601	0.765	0.711	0.759	0.700	0.201	0.341
	All, $ C \geq 10$	0.617	0.783	0.710	0.787	0.734	0.198	0.334
	Top 5000	0.499	0.635	0.622	0.629	0.565	0.190	0.294
	Top 5000, $ C \geq 10$	0.588	0.766	0.688	0.766	0.698	0.202	0.322
LiveJournal	All	0.563	0.875	0.589	0.874	0.774	0.336	0.489
	All, $ C \geq 10$	0.568	0.858	0.591	0.858	0.786	0.339	0.474
	Top 5000	0.150	0.295	0.163	0.295	0.430	0.101	0.100
	Top 5000, $ C \geq 10$	0.125	0.314	0.129	0.318	0.513	0.081	0.078
Orkut	All	0.718	0.916	0.731	0.917	0.928	0.750	0.711
	All, $ C \geq 10$	0.708	0.918	0.716	0.921	0.930	0.735	0.703
	Top 5000	0.550	0.863	0.554	0.862	0.891	0.584	0.564
	Top 5000, $ C \geq 10$	0.543	0.887	0.546	0.889	0.913	0.585	0.566
CYC/Gavin 2006		0.614	0.734	0.611	0.732	0.735	0.532	0.500
CYC/Krogan 2006		0.466	0.626	0.469	0.620	0.617	0.325	0.265
CYC/Collins 2007		0.716	0.953	0.712	0.953	0.972	0.720	0.730
CYC/Costanzo 2010		0.759	0.931	0.755	0.929	0.934	0.672	0.646
CYC/Hoppins 2011		0.788	0.883	0.785	0.882	0.970	0.763	0.763
CYC/all		0.674	0.872	0.742	0.874	0.840	0.363	0.026

Table 5: Standard deviation of conductance of the recovered communities with 1 seed.

Dataset	Communities	PGDC-0	PGDC-d	EMc-0	EMc-d	YL	HK	PPR
LFR (om=1)		0.018	0.159	0.031	0.162	0.098	0.018	0.015
LFR (om=2)		0.045	0.077	0.030	0.074	0.062	0.007	0.006
LFR (om=3)		0.037	0.074	0.026	0.071	0.060	0.006	0.005
LFR (om=4)		0.039	0.074	0.027	0.071	0.061	0.007	0.005
Karate		0.104	0.264	0.066	0.268	0.162	0.100	0.035
Football		0.061	0.139	0.068	0.148	0.140	0.038	0.042
Pol.Blogs		0.333	0.329	0.327	0.336	0.127	0.314	0.344
Pol.Books		0.155	0.210	0.122	0.221	0.168	0.089	0.057
Flickr		0.181	0.149	0.182	0.149	0.097	0.169	0.235
Amazon	All	0.110	0.254	0.113	0.254	0.241	0.052	0.025
	All, $ C \geq 10$	0.098	0.274	0.105	0.274	0.254	0.049	0.029
	Top 5000	0.061	0.244	0.069	0.243	0.222	0.033	0.021
	Top 5000, $ C \geq 10$	0.069	0.282	0.067	0.284	0.279	0.017	0.016
DBLP	All	0.104	0.257	0.111	0.254	0.229	0.066	0.056
	All, $ C \geq 10$	0.102	0.249	0.114	0.249	0.229	0.069	0.055
	Top 5000	0.096	0.257	0.109	0.252	0.228	0.056	0.056
	Top 5000, $ C \geq 10$	0.097	0.259	0.103	0.257	0.233	0.056	0.056
Youtube	All	0.182	0.191	0.191	0.189	0.195	0.097	0.123
	All, $ C \geq 10$	0.172	0.183	0.181	0.171	0.181	0.099	0.125
	Top 5000	0.204	0.220	0.229	0.220	0.208	0.096	0.124
	Top 5000, $ C \geq 10$	0.176	0.190	0.189	0.185	0.190	0.093	0.127
LiveJournal	All	0.179	0.170	0.195	0.168	0.151	0.156	0.189
	All, $ C \geq 10$	0.202	0.220	0.219	0.219	0.166	0.178	0.209
	Top 5000	0.174	0.339	0.195	0.338	0.347	0.121	0.143
	Top 5000, $ C \geq 10$	0.159	0.377	0.170	0.379	0.367	0.110	0.131
Orkut	All	0.170	0.124	0.183	0.123	0.039	0.164	0.192
	All, $ C \geq 10$	0.186	0.123	0.194	0.120	0.034	0.179	0.196
	Top 5000	0.157	0.182	0.165	0.184	0.108	0.167	0.167
	Top 5000, $ C \geq 10$	0.157	0.167	0.162	0.166	0.064	0.170	0.171
CYC/Gavin 2006		0.398	0.332	0.401	0.333	0.325	0.469	0.500
CYC/Krogan 2006		0.338	0.304	0.339	0.307	0.296	0.408	0.441
CYC/Collins 2007		0.449	0.161	0.453	0.161	0.055	0.443	0.432
CYC/Costanzo 2010		0.195	0.106	0.197	0.109	0.064	0.262	0.273
CYC/Hoppins 2011		0.385	0.285	0.388	0.285	0.068	0.425	0.425
CYC/all		0.084	0.118	0.116	0.116	0.088	0.038	0.093

Table 6: Average runtime for finding a community from 1 seed, in seconds.

Dataset	Communities	PGDC-0	PGDC-d	EMC-0	EMC-d	YL	HK	PPR
LFR (om=1)		0.056	0.548	0.035	0.604	0.069	0.034	0.649
LFR (om=2)		0.104	0.636	0.069	0.604	0.072	0.035	0.719
LFR (om=3)		0.106	0.641	0.077	0.615	0.070	0.035	0.735
LFR (om=4)		0.113	0.606	0.079	0.611	0.112	0.035	0.729
Karate		0.021	0.292	0.022	0.392	0.009	0.007	0.017
Football		0.040	0.406	0.021	0.376	0.019	0.006	0.046
Pol.Blogs		0.113	0.606	0.099	0.579	0.055	0.015	0.365
Pol.Books		0.067	0.608	0.024	0.407	0.011	0.006	0.040
Flickr		1.145	15.198	0.752	14.404	25.639	0.045	0.133
Amazon	All	0.068	0.831	0.062	0.963	4.418	0.031	2.803
	All, $ C \geq 10$	0.072	0.838	0.063	0.994	4.011	0.029	2.361
	Top 5000	0.049	0.804	0.051	1.004	2.629	0.022	1.643
	Top 5000, $ C \geq 10$	0.056	0.867	0.052	1.006	2.459	0.020	1.148
DBLP	All	0.082	0.906	0.070	1.100	5.799	0.039	1.405
	All, $ C \geq 10$	0.089	0.914	0.073	1.083	5.827	0.043	1.422
	Top 5000	0.070	0.859	0.061	1.025	5.637	0.036	1.319
	Top 5000, $ C \geq 10$	0.070	0.867	0.059	1.050	6.426	0.039	1.339
Youtube	All	0.304	3.892	0.310	6.295	18.975	0.074	0.514
	All, $ C \geq 10$	0.321	3.868	0.306	6.239	19.040	0.076	0.556
	Top 5000	0.241	3.815	0.313	6.314	18.295	0.078	0.520
	Top 5000, $ C \geq 10$	0.313	3.728	0.320	6.127	20.980	0.077	0.578
LiveJournal	All	0.422	7.444	0.504	10.008	189.899	0.115	0.312
	All, $ C \geq 10$	0.415	7.247	0.497	9.774	182.584	0.114	0.307
	Top 5000	0.400	7.204	0.512	10.506	101.696	0.101	0.446
	Top 5000, $ C \geq 10$	0.541	9.874	0.457	9.761	89.177	0.098	0.465
Orkut	All	3.703	70.340	3.471	69.008	1632.087	0.212	0.488
	All, $ C \geq 10$	3.787	71.671	3.469	70.452	1631.802	0.214	0.512
	Top 5000	1.521	27.182	1.253	26.215	1542.724	0.228	0.602
	Top 5000, $ C \geq 10$	1.404	25.700	1.244	25.161	1538.145	0.237	0.678
CYC/Gavin 2006		0.046	0.575	0.022	0.356	0.043	0.026	0.246
CYC/Krogan 2006		0.080	0.706	0.034	0.383	0.075	0.042	0.421
CYC/Collins 2007		0.052	0.475	0.016	0.307	0.042	0.024	0.249
CYC/Costanzo 2010		0.118	0.698	0.062	0.462	0.166	0.053	0.927
CYC/Hoppins 2011		0.038	0.425	0.015	0.292	0.036	0.024	0.160
CYC/all		0.189	0.982	0.071	0.589	0.386	0.052	1.760

Table 7: F_1 score with 3 seeds.

Dataset	Communities	PGDC-0	PGDC-d	EMC-0	EMC-d	YL	HK	PPR
LFR (om=1)		0.997	0.339	0.906	0.339	0.451	0.040	0.040
LFR (om=2)		0.583	0.217	0.277	0.204	0.283	0.039	0.041
LFR (om=3)		0.331	0.186	0.153	0.184	0.234	0.038	0.040
LFR (om=4)		0.227	0.176	0.108	0.173	0.212	0.036	0.038
Karate		0.855	0.611	0.811	0.612	0.832	0.888	0.957
Football		0.824	0.842	0.663	0.839	0.906	0.488	0.275
Pol.Blogs		0.793	0.048	0.793	0.053	0.026	0.776	0.616
Pol.Books		0.704	0.336	0.692	0.336	0.379	0.674	0.691
Flickr		0.114	0.040	0.114	0.040	0.017	0.053	0.150
Amazon	All	0.547	0.732	0.510	0.731	0.689	0.248	0.132
	Top 5000	0.944	0.801	0.944	0.800	0.831	0.730	0.538
	Top 5000, $ C \geq 10$	0.905	0.590	0.908	0.589	0.656	0.849	0.756
DBLP	All	0.399	0.550	0.360	0.550	0.539	0.223	0.219
	Top 5000	0.622	0.692	0.590	0.692	0.658	0.369	0.357
	Top 5000, $ C \geq 10$	0.555	0.520	0.516	0.520	0.490	0.385	0.385
Youtube	All	0.159	0.393	0.143	0.390	0.462	0.041	0.081
	Top 5000	0.368	0.556	0.338	0.555	0.545	0.097	0.164
	Top 5000, $ C \geq 10$	0.104	0.204	0.092	0.201	0.220	0.025	0.059
LiveJournal	All	0.098	0.545	0.085	0.545	0.317	0.040	0.065
	Top 5000	0.753	0.684	0.738	0.683	0.611	0.591	0.503
	Top 5000, $ C \geq 10$	0.700	0.580	0.684	0.577	0.513	0.577	0.467
Orkut	All	0.081	0.661	0.075	0.662	0.301	0.073	0.041
	Top 5000	0.421	0.235	0.412	0.235	0.185	0.376	0.313
	Top 5000, $ C \geq 10$	0.389	0.165	0.379	0.164	0.129	0.365	0.281
CYC/Gavin 2006		0.646	0.774	0.625	0.777	0.768	0.286	0.233
CYC/Krogan 2006		0.520	0.705	0.466	0.701	0.695	0.169	0.107
CYC/Collins 2007		0.502	0.730	0.501	0.730	0.692	0.284	0.284
CYC/Costanzo 2010		0.169	0.605	0.165	0.603	0.487	0.096	0.096
CYC/Hoppins 2011		0.495	0.624	0.493	0.621	0.691	0.282	0.282
CYC/all		0.298	0.801	0.253	0.799	0.505	0.021	0.002

Table 8: Mean community sizes with 3 seeds.

Dataset	Communities	PGDC-0	PGDC-d	EMC-0	EMC-d	YL	HK	PPR
LFR (om=1)		49.6	10.1	65.7	10.0	14.0	2418.7	2378.4
LFR (om=2)		115.7	6.9	354.3	6.6	10.5	2464.2	2316.9
LFR (om=3)		166.2	6.5	541.2	6.4	10.1	2457.1	2288.3
LFR (om=4)		200.1	6.6	630.7	6.5	10.2	2442.1	2268.0
Karate		22.0	8.5	24.7	8.6	13.8	17.2	17.3
Football		20.0	8.6	27.4	8.7	10.0	39.5	60.0
Pol.Blogs		626.1	32.4	648.1	36.9	10.4	573.0	1210.3
Pol.Books		48.4	9.0	53.2	9.2	10.2	49.4	54.7
Flickr		602.9	70.7	603.7	71.1	22.9	158.3	1210.5
Amazon	All	20.7	6.1	30.1	6.0	6.7	84.4	20781.4
	Top 5000	12.3	6.8	13.0	6.8	7.8	24.5	4855.6
	Top 5000, $ C \geq 10$	21.8	9.2	22.7	9.2	11.2	31.2	937.0
DBLP	All	85.0	6.9	113.1	6.9	6.6	52.9	25125.5
	Top 5000	26.5	6.6	33.1	6.6	6.6	44.0	8554.6
	Top 5000, $ C \geq 10$	44.0	8.2	57.4	8.1	7.1	48.9	7538.6
Youtube	All	214.5	20.4	269.0	20.9	11.8	124.0	19080.1
	Top 5000	117.6	11.2	169.2	12.2	8.0	102.2	11009.4
	Top 5000, $ C \geq 10$	341.4	17.2	449.0	18.3	11.4	142.4	20712.3
LiveJournal	All	196.0	6.3	221.2	6.4	19.0	155.5	3333.9
	Top 5000	56.4	21.5	60.3	21.3	13.7	100.2	528.6
	Top 5000, $ C \geq 10$	84.1	27.1	89.7	26.8	16.5	120.7	671.0
Orkut	All	192.7	13.6	201.6	13.7	18.7	214.8	1699.3
	Top 5000	289.4	45.7	295.2	45.0	18.6	277.5	1566.3
	Top 5000, $ C \geq 10$	306.9	42.9	315.8	42.7	20.3	291.9	1684.0
CYC/Gavin 2006		23.1	6.2	36.2	5.7	4.9	288.3	810.3
CYC/Krogan 2006		45.3	7.2	119.1	8.1	5.3	871.9	2063.1
CYC/Collins 2007		303.8	15.7	305.4	19.1	5.6	275.3	275.3
CYC/Costanzo 2010		713.5	39.4	738.4	41.5	10.4	1407.4	1240.4
CYC/Hoppins 2011		422.8	184.2	429.8	186.8	7.6	539.6	539.6
CYC/all		155.0	6.5	198.0	6.5	11.9	2636.0	5822.3

Table 9: Mean conductance of the recovered communities with 3 seeds.

Dataset	Communities	PGDC-0	PGDC-d	EMC-0	EMC-d	YL	HK	PPR
LFR (om=1)		0.300	0.730	0.305	0.730	0.644	0.245	0.268
LFR (om=2)		0.519	0.814	0.540	0.820	0.760	0.312	0.337
LFR (om=3)		0.580	0.824	0.575	0.825	0.776	0.333	0.353
LFR (om=4)		0.593	0.823	0.582	0.826	0.783	0.341	0.358
Karate		0.094	0.436	0.073	0.436	0.215	0.182	0.130
Football		0.265	0.415	0.232	0.411	0.353	0.266	0.160
Pol.Blogs		0.086	0.853	0.079	0.849	0.877	0.114	0.005
Pol.Books		0.089	0.578	0.069	0.575	0.565	0.139	0.066
Flickr		0.795	0.901	0.795	0.901	0.957	0.872	0.774
Amazon	All	0.189	0.477	0.188	0.479	0.467	0.080	0.054
	Top 5000	0.062	0.234	0.063	0.236	0.197	0.039	0.030
	Top 5000, $ C \geq 10$	0.039	0.366	0.042	0.368	0.317	0.020	0.018
DBLP	All	0.246	0.613	0.253	0.612	0.646	0.127	0.145
	Top 5000	0.193	0.472	0.197	0.473	0.507	0.109	0.118
	Top 5000, $ C \geq 10$	0.191	0.533	0.197	0.536	0.611	0.104	0.115
Youtube	All	0.606	0.792	0.656	0.789	0.819	0.201	0.337
	Top 5000	0.484	0.663	0.547	0.661	0.666	0.195	0.292
	Top 5000, $ C \geq 10$	0.626	0.831	0.703	0.824	0.845	0.196	0.316
LiveJournal	All	0.561	0.912	0.568	0.912	0.804	0.337	0.489
	Top 5000	0.144	0.364	0.145	0.365	0.456	0.100	0.092
	Top 5000, $ C \geq 10$	0.116	0.425	0.117	0.428	0.537	0.078	0.070
Orkut	All	0.700	0.935	0.700	0.935	0.919	0.732	0.704
	Top 5000	0.518	0.809	0.517	0.811	0.873	0.556	0.546
	Top 5000, $ C \geq 10$	0.517	0.831	0.517	0.832	0.890	0.562	0.552
CYC/Gavin 2006		0.507	0.638	0.510	0.641	0.655	0.400	0.353
CYC/Krogan 2006		0.382	0.546	0.396	0.547	0.565	0.193	0.123
CYC/Collins 2007		0.578	0.946	0.576	0.943	0.956	0.591	0.605
CYC/Costanzo 2010		0.700	0.927	0.695	0.927	0.915	0.601	0.538
CYC/Hoppins 2011		0.613	0.799	0.608	0.797	0.945	0.567	0.567
CYC/all		0.698	0.902	0.701	0.902	0.815	0.368	0.022

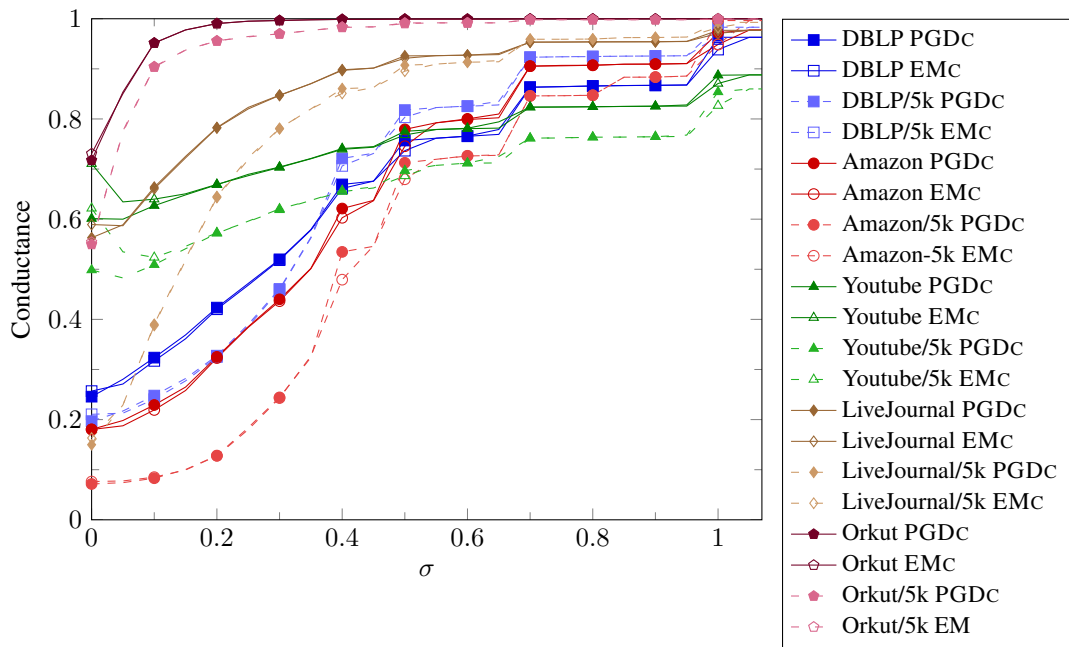


Figure 3: Conductance as a function of the σ parameter on the SNAP datasets with 1 seed.