Random Generation of Shortest Path Test Networks

Dennis J. Adams-Smith • Douglas R. Shier

Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-0975, USA

dadamss@clemson.edu • *shierd@clemson.edu*

The exact analytic results for graphs with $n = 4$ nodes and $m = 8$ edges are based on the following 16 spanning trees $T_1, T_2, \ldots, T_{16}$, all rooted at node $r = 1$:

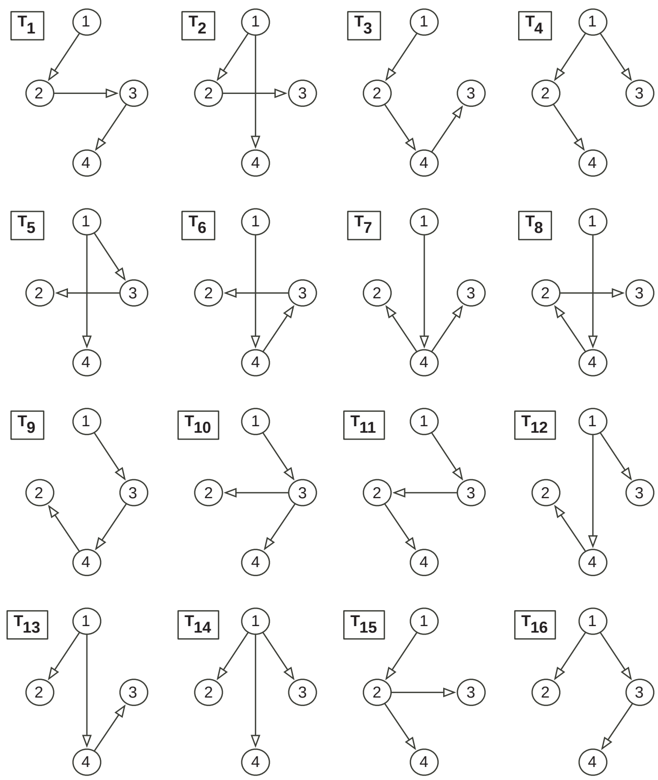

Figure 1: The 16 possible 1-rooted spanning trees on 4 nodes

The following sections tabulate computational results for generating root-connected directed networks G, with n nodes and m edges, using the following seven methods:

Uniform uses the G_{nm} random model and checks for root-connectivity

Prim generates the initial spanning tree T using Prim's algorithm

Kruskal generates the initial spanning tree T using Kruskal's algorithm

Broder generates the initial spanning tree T using Broder's random walk

Wilson generates the initial spanning tree T using Wilson's random walk

Hamiltonian Cycle generates an initial random Hamiltonian cycle

One-In generates random predecessors for all nodes other than the root

Random networks G were generated for each network size (n, m), with $R = 200$ replications at each size. In the tables that follow, we report the mean values for

Deg(r) the outdegree of the root node r in G

$D_2(G)$ the number of nodes at distance two from r in G

Diam the maximum length of a shortest path (number of edges) from r to a node of G

Standard deviations are also reported for the first two measures.

1 Tables for $n = 100$

In this section we present results for networks with 100 nodes and 250, 300, 350, 400, 450, 500, 600 edges in Tables 1–7 respectively. Here the network density $\delta = \frac{m}{n}$ has the range $2.5 \leq \delta \leq 6$. In general, the `One-In` and `Hamiltonian Cycle` methods give values closest to those for `Uniform`. The random spanning tree methods (`Broder`, `Wilson`) and `Kruskal` give values very close to one other, but more biased than the `One-In` and `Hamiltonian Cycle` methods. The `Prim` method is the most biased, and this bias persists at higher densities, when one might have expected it to become less evident.

Table 1: Computational results for networks with 100 nodes, 250 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	2.835	1.578	6.855	4.315	8.620
One-In	2.775	1.416	6.500	3.980	8.210
Hamiltonian Cycle	2.510	1.182	5.990	3.320	8.210
Kruskal	3.595	1.632	9.160	4.586	7.555
Broder	3.365	1.614	8.525	4.265	7.635
Wilson	3.315	1.475	8.780	4.338	7.630
Prim	6.420	2.345	20.505	6.085	5.770

Table 2: Computational results for networks with 100 nodes, 300 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	3.050	1.674	8.555	4.759	7.230
One-In	3.360	1.650	8.960	4.618	6.705
Hamiltonian Cycle	3.040	1.363	8.595	3.883	6.955
Kruskal	3.870	1.714	11.670	4.910	6.475
Broder	4.065	1.799	11.360	5.026	6.545
Wilson	4.005	1.618	11.165	5.051	6.595
Prim	6.995	2.411	23.965	6.796	5.140

Table 3: Computational results for networks with 100 nodes, 350 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	3.495	1.635	11.330	5.554	6.260
One-In	3.720	1.934	11.880	6.227	6.060
Hamiltonian Cycle	3.350	1.549	10.920	5.193	6.120
Kruskal	4.385	1.747	14.350	5.934	5.800
Broder	4.615	1.812	14.825	6.070	5.690
Wilson	4.735	1.877	15.555	6.353	5.720
Prim	7.525	2.430	27.330	7.355	4.815

Table 4: Computational results for networks with 100 nodes, 400 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	4.175	1.814	15.020	7.015	5.675
One-In	4.025	1.877	14.435	6.693	5.480
Hamiltonian Cycle	4.005	1.676	14.195	5.635	5.460
Kruskal	5.070	2.254	18.360	7.431	5.170
Broder	4.890	1.851	17.375	6.334	5.245
Wilson	5.245	1.999	18.800	6.188	5.180
Prim	8.010	2.490	31.335	7.861	4.425

Table 5: Computational results for networks with 100 nodes, 450 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	4.445	2.071	17.065	7.329	5.230
One-In	4.520	2.032	17.730	7.798	5.015
Hamiltonian Cycle	4.525	1.837	17.415	7.060	5.060
Kruskal	5.335	1.981	21.435	7.490	4.820
Broder	5.185	2.206	20.465	8.128	4.865
Wilson	5.660	1.961	22.045	6.709	4.895
Prim	8.525	2.624	34.460	8.097	4.285

Table 6: Computational results for networks with 100 nodes, 500 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	5.010	2.025	21.250	8.378	4.810
One-In	5.045	2.096	20.865	7.850	4.705
Hamiltonian Cycle	5.090	1.939	21.155	7.152	4.700
Kruskal	5.820	2.207	24.630	8.175	4.520
Broder	5.835	2.037	24.510	7.755	4.510
Wilson	5.835	1.982	24.540	7.387	4.460
Prim	9.095	2.689	37.810	7.863	4.045

Table 7: Computational results for networks with 100 nodes, 600 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	6.155	2.478	28.490	9.442	4.235
One-In	5.960	2.380	28.155	10.071	4.165
Hamiltonian Cycle	6.165	2.191	28.790	8.979	4.140
Kruskal	6.725	2.462	31.595	9.297	4.095
Broder	6.920	2.245	32.735	8.397	4.070
Wilson	6.845	2.370	31.865	9.135	4.115
Prim	9.590	2.507	44.390	7.744	3.775

2 Tables for $n = 200$

In this section we present results for networks with 200 nodes and 600, 650, 800, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000 edges in Tables 8–19 respectively. Here the network density δ has the rather large range $3 \leq \delta \leq 25$. Again, the `One-In` and `Hamiltonian Cycle` methods give values closest to those for `Uniform`. The random spanning tree methods (`Broder`, `Wilson`) and `Kruskal` give similar values but they are more biased than the `One-In` and `Hamiltonian Cycle` methods, especially at lower network densities. The bias in the `Prim` method persists even at the relatively high density level of $\delta = 25$.

Table 8: Computational results for networks with 200 nodes, 600 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	3.110	1.575	8.875	5.066	8.240
One-In	3.215	1.662	8.940	5.144	8.055
Hamiltonian Cycle	3.120	1.313	9.330	4.341	7.960
Kruskal	4.035	1.746	12.140	5.533	7.575
Broder	3.950	1.840	12.290	5.978	7.610
Wilson	4.020	1.701	12.800	5.348	7.425
Prim	7.785	2.496	30.000	8.903	5.985

Table 9: Computational results for networks with 200 nodes, 650 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	3.430	1.844	10.900	6.344	7.745
One-In	3.300	1.675	10.250	5.802	7.480
Hamiltonian Cycle	3.535	1.403	11.150	4.860	7.330
Kruskal	4.140	1.629	13.520	5.786	7.010
Broder	4.340	1.822	14.485	6.352	7.015
Wilson	4.080	1.705	13.950	6.124	7.030
Prim	7.960	2.575	32.790	10.361	5.720

Table 10: Computational results for networks with 200 nodes, 800 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	4.195	1.946	15.955	8.001	6.465
One-In	3.995	1.888	15.300	7.855	6.335
Hamiltonian Cycle	3.860	1.601	14.490	6.458	6.350
Kruskal	4.980	2.093	19.810	8.279	5.955
Broder	5.025	2.092	20.040	8.339	5.925
Wilson	4.950	1.951	19.595	7.979	6.020
Prim	8.890	2.710	40.900	10.803	5.175

Table 11: Computational results for networks with 200 nodes, 1000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	5.045	2.251	22.975	10.485	5.520
One-In	5.100	2.178	23.875	10.260	5.335
Hamiltonian Cycle	5.000	2.253	22.975	10.056	5.320
Kruskal	6.020	2.103	28.455	9.455	5.100
Broder	6.240	2.287	29.190	10.502	5.125
Wilson	5.980	2.166	27.765	9.381	5.165
Prim	9.985	2.801	51.390	12.321	4.570

Table 12: Computational results for networks with 200 nodes, 1500 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	7.515	2.612	47.320	14.625	4.230
One-In	7.475	2.631	46.530	15.185	4.145
Hamiltonian Cycle	7.545	2.631	47.225	14.984	4.145
Kruskal	8.325	2.709	52.380	14.969	4.075
Broder	8.245	2.475	51.710	13.480	4.110
Wilson	8.190	2.537	51.390	13.130	4.080
Prim	12.865	3.217	78.575	14.746	3.915

Table 13: Computational results for networks with 200 nodes, 2000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	10.345	3.029	76.420	16.125	3.780
One-In	9.920	3.121	74.380	17.326	3.750
Hamiltonian Cycle	10.135	3.122	74.630	17.372	3.795
Kruskal	11.185	3.178	81.750	17.612	3.610
Broder	10.780	3.003	79.075	16.822	3.735
Wilson	10.870	3.013	79.755	16.222	3.685
Prim	14.835	3.546	102.555	15.552	3.285

Table 14: Computational results for networks with 200 nodes, 2500 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	12.555	3.356	101.170	17.899	3.205
One-In	12.880	3.474	102.405	17.839	3.185
Hamiltonian Cycle	12.160	3.229	99.710	17.135	3.160
Kruskal	12.750	3.474	103.030	17.339	3.160
Broder	13.405	3.290	106.425	17.293	3.100
Wilson	13.540	3.099	107.350	15.226	3.160
Prim	16.795	3.899	122.270	14.305	3.040

Table 15: Computational results for networks with 200 nodes, 3000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	14.745	3.754	123.250	15.626	3.015
One-In	14.945	3.791	124.200	15.854	3.010
Hamiltonian Cycle	15.075	3.251	124.925	13.988	3.005
Kruskal	16.155	3.592	129.325	13.794	3.010
Broder	15.790	3.690	128.225	15.518	3.010
Wilson	15.990	3.626	128.355	15.148	3.005
Prim	19.785	3.939	141.840	10.446	3.000

Table 16: Computational results for networks with 200 nodes, 3500 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	17.425	3.839	142.375	12.512	3.000
One-In	17.165	4.188	141.085	15.633	3.005
Hamiltonian Cycle	17.395	3.951	142.740	12.394	3.000
Kruskal	18.340	3.685	145.850	9.294	3.000
Broder	18.940	3.643	146.345	10.369	3.000
Wilson	18.400	4.020	145.420	11.476	3.000
Prim	21.750	4.088	152.460	7.368	3.000

Table 17: Computational results for networks with 200 nodes, 4000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	19.970	4.286	154.425	8.421	3.000
One-In	19.810	4.509	154.660	8.726	3.000
Hamiltonian Cycle	19.945	4.251	154.835	8.478	3.000
Kruskal	20.845	3.982	156.985	6.825	3.000
Broder	20.440	3.922	157.130	6.940	3.000
Wilson	21.100	4.296	157.310	7.420	3.000
Prim	24.105	4.669	160.665	5.521	3.000

Table 18: Computational results for networks with 200 nodes, 4500 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	22.300	4.164	162.825	6.215	3.000
One-In	23.065	4.995	162.510	6.096	2.995
Hamiltonian Cycle	22.350	4.299	162.345	5.793	3.000
Kruskal	22.915	4.352	162.490	5.270	2.995
Broder	23.420	4.377	163.825	4.224	3.000
Wilson	23.555	4.053	163.825	4.443	3.000
Prim	26.925	4.798	164.440	3.491	2.970

Table 19: Computational results for networks with 200 nodes, 5000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	25.145	4.802	166.645	3.594	2.975
One-In	25.185	4.985	166.560	3.349	2.975
Hamiltonian Cycle	25.055	4.359	166.680	3.194	2.965
Kruskal	25.595	4.930	166.435	3.430	2.955
Broder	26.025	4.494	166.785	3.130	2.980
Wilson	26.245	4.302	166.935	2.681	2.970
Prim	29.085	4.701	166.165	3.173	2.895

3 Tables for $n = 400$

In this section we present results for networks with 400 nodes and 2000, 2500, 3000, 4000, 5000, 6000 edges in Tables 20–25 respectively. Here the network density δ has the range $5 \leq \delta \leq 15$. Again, the `One-In` and `Hamiltonian Cycle` methods give values closest to those for `Uniform`. The random spanning tree methods (`Broder`, `Wilson`) and `Kruskal` give similar values but they are more biased than the `One-In` and `Hamiltonian Cycle` methods, especially at lower network densities. The bias in the `Prim` method persists at every density level.

Table 20: Computational results for networks with 400 nodes, 2000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	5.005	2.019	23.480	10.726	6.275
One-In	4.765	2.266	22.645	11.190	6.110
Hamiltonian Cycle	4.900	2.008	23.820	10.575	6.005
Kruskal	5.990	2.235	30.350	11.312	5.775
Broder	5.960	2.250	29.095	11.933	5.865
Wilson	5.810	2.111	28.835	10.837	5.810
Prim	10.515	2.851	61.265	14.918	5.135

Table 21: Computational results for networks with 400 nodes, 2500 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	6.450	2.500	37.660	15.201	5.295
One-In	6.140	2.452	36.560	14.973	5.160
Hamiltonian Cycle	6.200	2.437	36.350	13.998	5.115
Kruskal	7.360	2.755	43.555	15.600	5.085
Broder	7.320	2.344	43.325	13.915	5.085
Wilson	7.125	2.420	42.315	14.473	5.070
Prim	11.600	3.186	76.345	18.330	4.595

Table 22: Computational results for networks with 400 nodes, 3000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	7.790	2.777	53.655	18.301	4.785
One-In	7.470	2.834	51.445	18.712	4.670
Hamiltonian Cycle	7.285	2.615	50.315	18.474	4.705
Kruskal	8.250	2.648	57.315	17.990	4.605
Broder	8.720	2.964	59.450	18.707	4.545
Wilson	8.545	2.692	58.200	17.984	4.555
Prim	12.895	3.183	95.005	20.202	4.135

Table 23: Computational results for networks with 400 nodes, 4000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	9.890	3.021	85.280	23.119	4.030
One-In	10.145	3.128	88.125	25.022	4.040
Hamiltonian Cycle	9.900	2.756	86.060	22.834	4.015
Kruskal	11.345	3.257	96.760	24.586	4.015
Broder	10.910	3.177	93.610	25.281	4.025
Wilson	10.810	3.160	92.260	24.261	4.020
Prim	15.145	3.570	131.585	25.188	3.975

Table 24: Computational results for networks with 400 nodes, 5000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	12.165	3.385	122.855	28.510	3.915
One-In	12.225	3.573	123.620	30.378	3.915
Hamiltonian Cycle	12.480	3.417	125.425	28.045	3.935
Kruskal	13.170	3.273	131.535	25.717	3.895
Broder	13.245	3.698	131.715	30.077	3.880
Wilson	13.420	3.491	132.315	27.264	3.855
Prim	17.560	3.624	169.585	24.549	3.490

Table 25: Computational results for networks with 400 nodes, 6000 edges, 200 trials

Method	Mean Deg(r)	STD Deg(r)	Mean $D_2(G)$	STD $D_2(G)$	Mean diam
Uniform	14.805	3.516	163.970	29.253	3.385
One-In	15.180	3.885	167.135	32.017	3.315
Hamiltonian Cycle	15.100	3.591	166.090	30.088	3.375
Kruskal	15.450	3.804	169.275	30.685	3.315
Broder	15.930	3.921	173.005	31.351	3.305
Wilson	15.860	3.705	172.340	29.484	3.295
Prim	20.335	3.978	209.780	27.150	3.075