Supplementary Information

Competition-Based Benchmarking of Influence Ranking Methods in Social Networks

Alexandru Topîrceanu

Contents

1 Node overlapping correlation change as the number of selected spreaders increases

Given the four synthetic networks (10,000 nodes), we alternatively measure the amount of node overlapping for 10 selected influence ranking methods. These methods are: node degree centrality (Deg), closeness (Cls)[1], node betweenness (Btw)[2], HITS[3], PageRank (PR)[4], Hirsch index (HI)[5], LeaderRank (LR)[6], K-shell decomposition (KS)[7], Local centrality with a coefficient (CLC)[8], and Eigenvector centrality (EC)[9].

The 10 mentioned centralities are applied in order to select the top $p\%$ spreaders in each network, for values of $p \in \{0.01, 0.05, 0.1\}$. Figure 1 shows the change in correlation (node overlapping) by increasing $p = 0.01$ to $p = 0.1$. The correlation between most centralities will drop as p increases. The average changes in correlations from $p = 0.01$ up to $p = 0.1$ are: $\delta_{Rand} = -0.289$, $\delta_{Mesh} = -0.193$, $\delta_{SW} = -0.189$, and $\delta_{SF} = -0.088$.

The upper panel of Figure 1 shows the correlation drop grouped by centrality measures, and the lower panel of Figure 1 groups the results by topology. We notice that the scale-free network has the lowest drop in correlation. This can be explained by the existence of hubs in the network, and the fact that most centralities will select those few nodes within the top $p\%$ spreader set. This is why, even if the spreader pool increases to 10% of the network, the same spreaders are selected, so the correlation remains high and mostly unaltered.

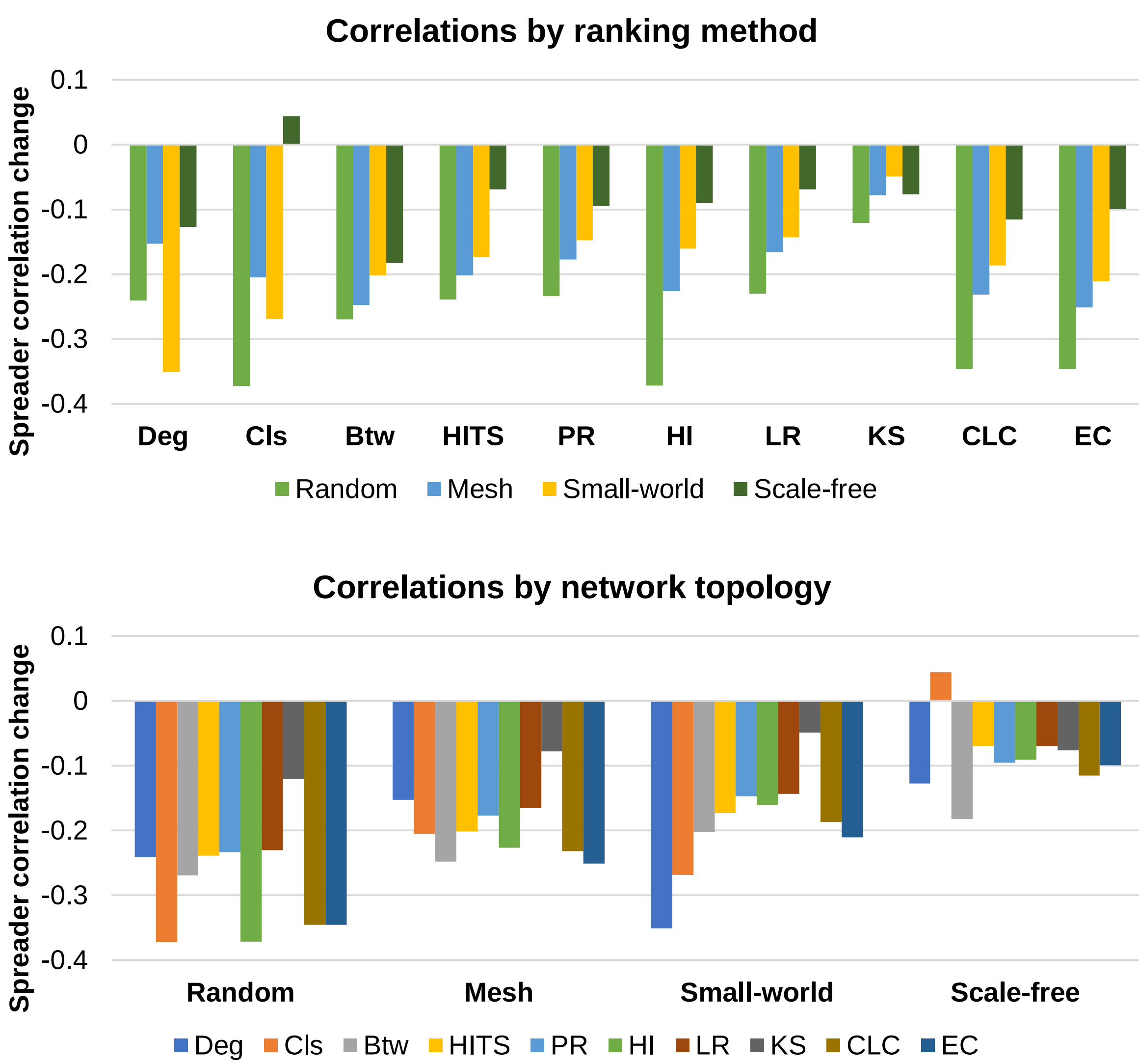

Figure 1. Changes in correlation (node overlapping) between ranking methods by increasing the spreader size from 1% to 10% of the total network size. Each synthetic network has $N = 10,000$ nodes.

On the other hand, the random network produces a significant drop in correlation. This is due to the fact that, as more nodes are introduced in the spreader pool, more randomness is introduced, so the overlapping drops naturally. In an analogous manner, the mesh network also has a high correlation drop, and this is due to the fact that influential nodes tend to have a uniform spatial distribution. Selecting a larger spreader size, incurs higher overall diversity, as in a Gaussian distributed population. Finally, the small-world network is similar to a mesh network with an additional set of long range links ($< 15\%$), so it performs very similar to the mesh topology.

We can conclude that only the scale-free property has a more characteristic response, due to its small subset of hub nodes, which naturally represents around $1-10\%$ of the network size.

2 Node overlapping in terms of selected spreaders

Figure 2 exemplifies the spatial distribution of four selected centralities (degree, closeness, betweenness and PageRank) on the mesh topology, as they are used to select the top 1% nodes as spreaders. With this example we highlight the change in diversity of spreader positioning between the various ranking methods. In the selected example, degree and PageRank have similar spatial distributions, while Closeness is consistently different.

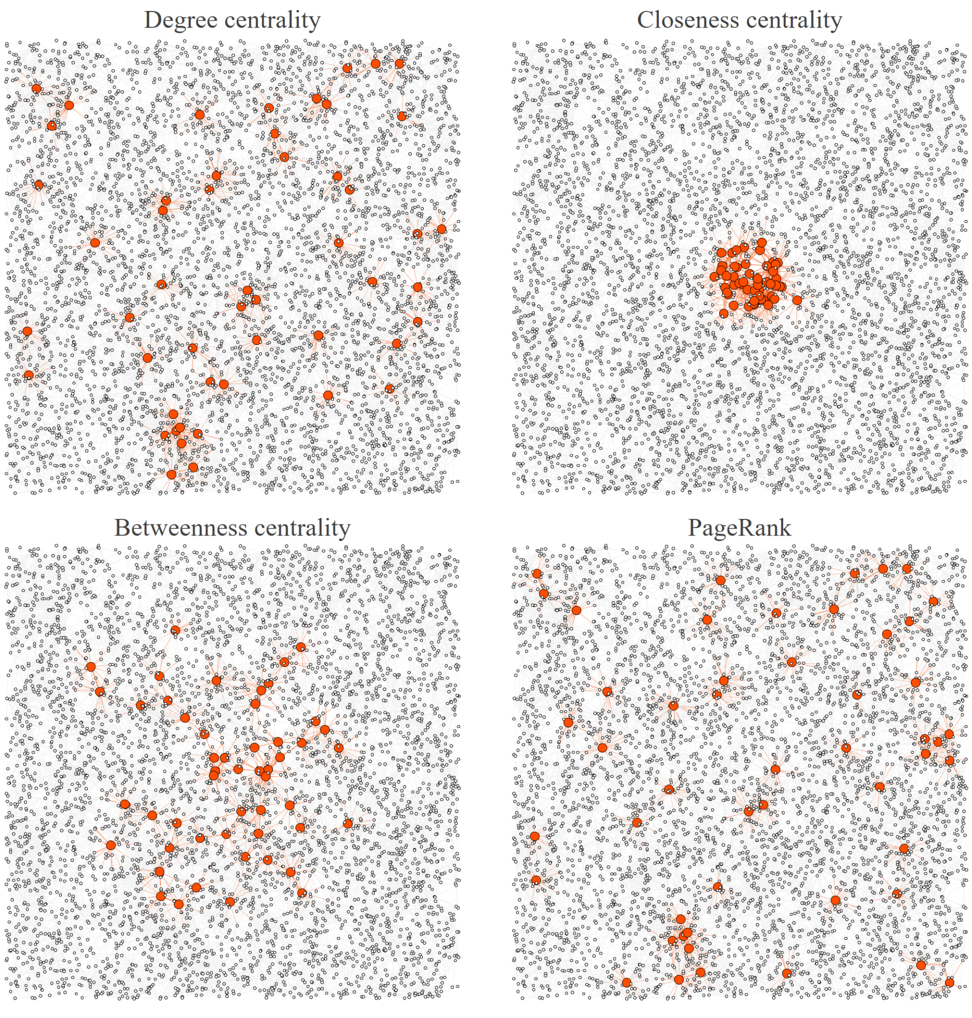

Figure 2. Spatial distribution of selected spreader nodes on the mesh network with $N = 10,000$ nodes. The top 1% nodes are highlighted as spreaders, as determined by the degree, closeness, betweenness and PageRank centralities.

3 Competition-based benchmarking results

Tables 1 and 2 contain the main detailed benchmarking simulation results of the paper. Each of the 8 datasets has one independent sub-table containing 10×10 averaged simulation sets between the following centrality measures: : node degree centrality (Deg), closeness (Cls), node betweenness (Btw), HITS, PageRank (PR), Hirsch index (HI), LeaderRank (LR), K-shell decomposition (KS), Local centrality with a coefficient (CLC), and Eigenvector centrality (EC).

Table 1. Synthetic dataset (random, mesh, small-world, scale-free) benchmark results for pair-wise competition between centrality measures. Each cell (x, y) contains the final opinion coverage $(0-100\%)$ for centrality x; the symmetric cell (y, x) represents the same number on a colour gradient blue (0%)–white (50%)–orange (100%).

Random	Deg	Cls	Btw	HITS	PR	HI	LR	KS	CLC	EC	*Avg*
Deg	-	98.50	1.82	0.98	1.00	98.90	98.10	99.02	98.78	98.56	66.18
Cls		-	1.30	1.48	1.50	98.12	1.74	98.76	1.30	1.36	23.02
Btw			-	1.30	1.60	98.80	97.84	98.82	98.52	98.60	66.15
HITS				-	1.38	98.84	1.24	99.10	98.78	98.62	66.28
PR					-	98.50	2.10	99.14	98.74	98.46	**77.16**
HI						-	1.78	98.36	1.18	1.00	12.13
LR							-	98.90	98.24	98.68	76.95
KS								-	0.98	0.82	0.99
CLC									-	98.86	33.93
EC										-	23.12

Mesh	Deg	Cls	Btw	HITS	PR	HI	LR	KS	CLC	EC	*Avg*
Deg	-	95.36	78.74	57.00	56.22	78.04	56.72	72.66	62.46	84.18	**71.26**
Cls		-	3.04	4.66	5.54	5.20	4.94	5.46	5.36	10.64	5.47
Btw			-	19.74	22.14	33.60	23.92	79.56	44.30	44.36	42.93
HITS				-	52.80	70.60	45.88	79.92	63.70	88.00	69.32
PR					-	65.52	44.56	71.00	67.66	73.60	65.35
HI						-	33.10	51.12	41.14	82.38	52.82
LR							-	77.54	62.46	79.60	67.57
KS								-	41.74	47.46	39.65
CLC									-	78.90	52.36
EC										-	32.96

SW	Deg	Cls	Btw	HITS	PR	HI	LR	KS	CLC	EC	*Avg*
Deg	-	89.42	66.94	60.16	34.58	78.56	66.60	70.36	79.36	74.44	68.94
Cls		-	9.10	5.82	12.82	11.58	16.26	23.06	8.02	8.24	11.39
Btw			-	21.24	47.48	43.40	44.88	79.66	59.80	84.14	56.96
HITS				-	54.06	90.58	71.60	86.02	52.20	93.02	**76.92**
PR					-	74.38	48.36	74.24	87.24	61.00	71.93
HI						-	51.38	83.24	39.54	31.06	33.25
LR							-	86.60	54.28	82.06	66.72
KS								-	47.10	25.32	37.87
CLC									-	41.96	60.24
EC										-	39.43

SF	Deg	Cls	Btw	HITS	PR	HI	LR	KS	CLC	EC	*Avg*
Deg	-	98.22	46.16	42.12	59.30	47.64	52.50	61.42	79.70	68.30	**61.71**
Cls		-	1.82	1.72	1.80	1.82	1.80	1.88	1.90	1.86	1.83
Btw			-	43.08	40.48	63.96	38.60	87.92	80.80	71.34	62.78
HITS				-	54.54	48.84	50.24	61.92	79.80	72.42	61.63
PR					-	52.96	31.20	52.44	80.84	67.76	55.74
HI						-	35.18	70.44	87.66	66.64	54.72
LR							-	62.88	79.66	72.84	61.53
KS								-	64.06	41.78	45.89
CLC									-	14.08	26.99
EC										-	43.09

Table 2. Real-world dataset benchmark results for pair-wise competition between centrality measures. Each cell (x,y) contains the final opinion coverage $(0-100\%)$ for centrality x; the symmetric cell (y,x) represents the same number on a colour gradient blue (0%)–white (50%)–orange (100%).

OSN	Deg	Cls	Btw	HITS	PR	HI	LR	KS	CLC	EC	*Avg*
Deg	-	97.52	78.35	8.16	81.99	92.20	7.68	88.31	11.79	8.84	52.76
Cls		-	1.79	1.31	2.21	3.94	1.42	2.73	4.42	2.52	2.55
Btw			-	13.05	12.42	90.36	13.58	90.52	15.79	15.90	40.37
HITS				-	82.67	89.46	10.26	91.67	89.67	17.58	**64.42**
PR					-	18.79	13.79	90.25	16.01	16.22	41.08
HI						-	7.53	7.16	4.31	5.63	24.23
LR							-	92.15	91.15	17.21	64.39
KS								-	11.37	12.11	28.77
CLC									-	12.48	44.74
EC										-	62.83

FB	Deg	Cls	Btw	HITS	PR	HI	LR	KS	CLC	EC	*Avg*
Deg	-	96.59	66.86	36.97	35.49	69.54	30.17	69.45	30.58	69.95	56.18
Cls		-	3.49	3.56	5.26	3.53	4.98	70.23	5.23	3.53	11.49
Btw			-	34.45	42.68	74.59	56.52	73.67	57.09	48.23	57.51
HITS				-	57.18	70.21	31.58	70.33	37.04	69.92	62.10
PR					-	70.21	39.53	70.20	43.85	69.89	55.99
HI						-	29.76	70.01	29.79	30.55	41.36
LR							-	70.23	64.50	68.89	**68.06**
KS								-	29.66	30.48	28.87
CLC									-	69.29	55.91
EC										-	44.54

Emails	Deg	Cls	Btw	HITS	PR	HI	LR	KS	CLC	EC	*Avg*
Deg	-	98.66	56.89	49.70	47.79	69.91	50.17	68.05	64.76	65.75	63.52
Cls		-	1.06	1.08	1.06	5.97	1.07	2.08	6.17	2.03	2.40
Btw			-	40.81	40.87	68.12	40.79	67.22	63.61	63.59	58.33
HITS				-	48.01	70.00	50.01	67.31	64.64	65.71	63.56
PR					-	69.07	51.14	67.66	63.75	64.00	63.55
HI						-	29.91	48.16	30.17	33.52	39.60
LR							-	71.41	65.18	65.04	**63.97**
KS								-	36.00	38.26	42.07
CLC									-	44.46	48.43
EC										-	51.49

POK	Deg	Cls	Btw	HITS	PR	HI	LR	KS	CLC	EC	*Avg*
Deg	-	71.00	50.12	55.34	41.23	80.42	35.59	89.36	73.23	73.21	63.28
Cls		-	38.60	29.16	31.01	53.54	28.27	84.58	53.74	62.13	45.78
Btw			-	47.22	39.32	59.18	50.48	92.35	68.10	58.49	58.27
HITS				-	45.48	70.45	37.91	93.26	71.85	80.42	63.09
PR					-	70.01	52.25	90.52	69.56	75.22	63.94
HI						-	21.17	85.66	30.62	45.73	36.30
LR							-	89.89	72.35	77.46	**66.87**
KS								-	14.11	19.11	13.33
CLC									-	79.53	48.01
EC										-	32.27

Each numerical cell (x,y) in the tables represents the opinion coverage $(0-100\%)$ for centrality x (versus y) at the end of each simulation, after balancing is attained[10] in our tolerance diffusion model[11]. The table cells are symmetrical, meaning that each colour cell (y,x) represents the same percentage on a colour gradient blue (0%)–white (50%)–orange (100%) For example, the values and colours are interpreted as follows: in the Emails dataset in Table 2, we obtain the result HI-LR=29.91, meaning that HI obtains 29.91% coverage in the network, while LR obtains the rest of $100-29.91 = 70.09\%$ coverage. Accordingly, cell LR-HI is coloured in a light blue, meaning that the first centrality wins (*i.e.*, by 70.09%). Also, the numerical values in each cell take into consideration AOA, meaning that simulations were automatically run in both scenarios where HI had priority in assigning spreaders, then LR had priority over HI.

Each cell in Tables 1-2 corresponds to a total of 10 repeated simulation batches (*i.e.*, translating into 20 because of AOA), thus a sub-table comprises a total of $45 \times 20 = 900$ simulations, amassing to an overall $8 \times 900 = 7,200$ unique simulations.

4 Ranking method performances on each dataset for individual SIR benchmarking

In Figure 3 we highlight the low numerical differentiation provided by testing influence ranking methods in an individual context with SIR epidemic diffusion[12,13]. The performance is expressed as percentage of final coverage at the end of a SIR epidemic diffusion. Overall, there is no visual cue suggesting that ranking methods perform any different on, *e.g.*, the SW topology (grey bars in Figure 3). In other words, we cannot reliably rank node centralities on a given topology. We can only rank the spreading ability when comparing different topologies, like for example, the Facebook dataset is covered to a larger extent than POK.

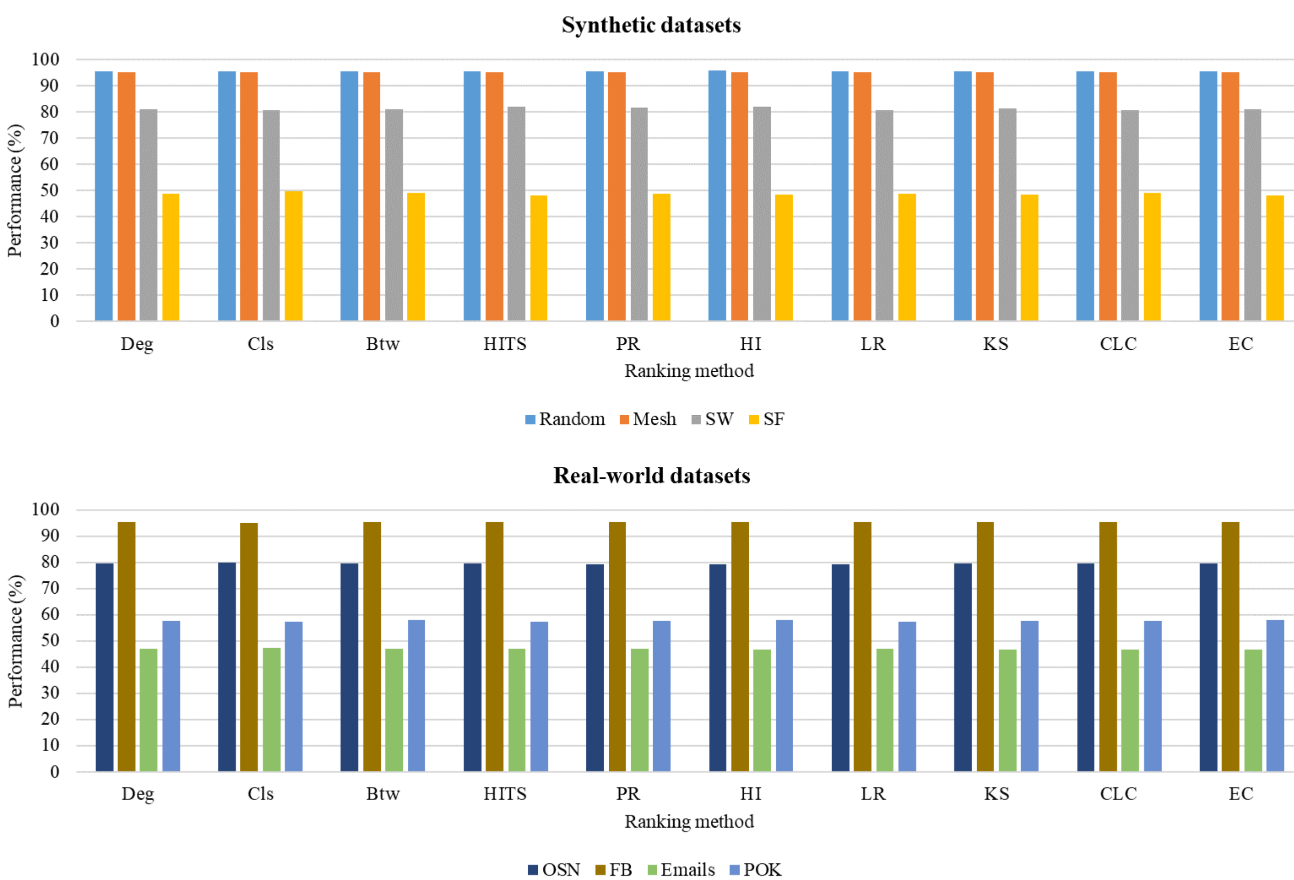

Figure 3. Performance (coverage 0-100%) of each ranking method on the 8 datasets using individual SIR benchmarking.

5 Ranking method performances on each dataset for competition-based benchmarking

In Figure 4 we display the averaged benchmark performances of each ranking method, for the 8 individual datasets. On the synthetic datasets, we notice that Deg, HITS, PR and LR have overall both stable and good results ($> 60\%$). On the other hand, Cls and KS have the weakest overall results ($< 30\%$). Additionally, some measures present significant change based on topology – EC increases in performance on scale-free networks; CLC is efficient only on mesh-like topologies; HI is promising on meshes, weak on random networks, and moderately performant on small-worlds. On the real-world datasets, we find high coverage stability for HITS and LR ($63-67\%$), which also come out as the best ranking methods overall. Cls is highly inefficient ($< 10\%$), except on the POK dataset (47%), while EC is generally efficient ($> 52\%$), except on the POK dataset (34%). These results represent complex emergent opinion coverages that take into consideration the composition of fundamental topologies found the real-world datasets, and their interpretation is beyond the scope of this paper.

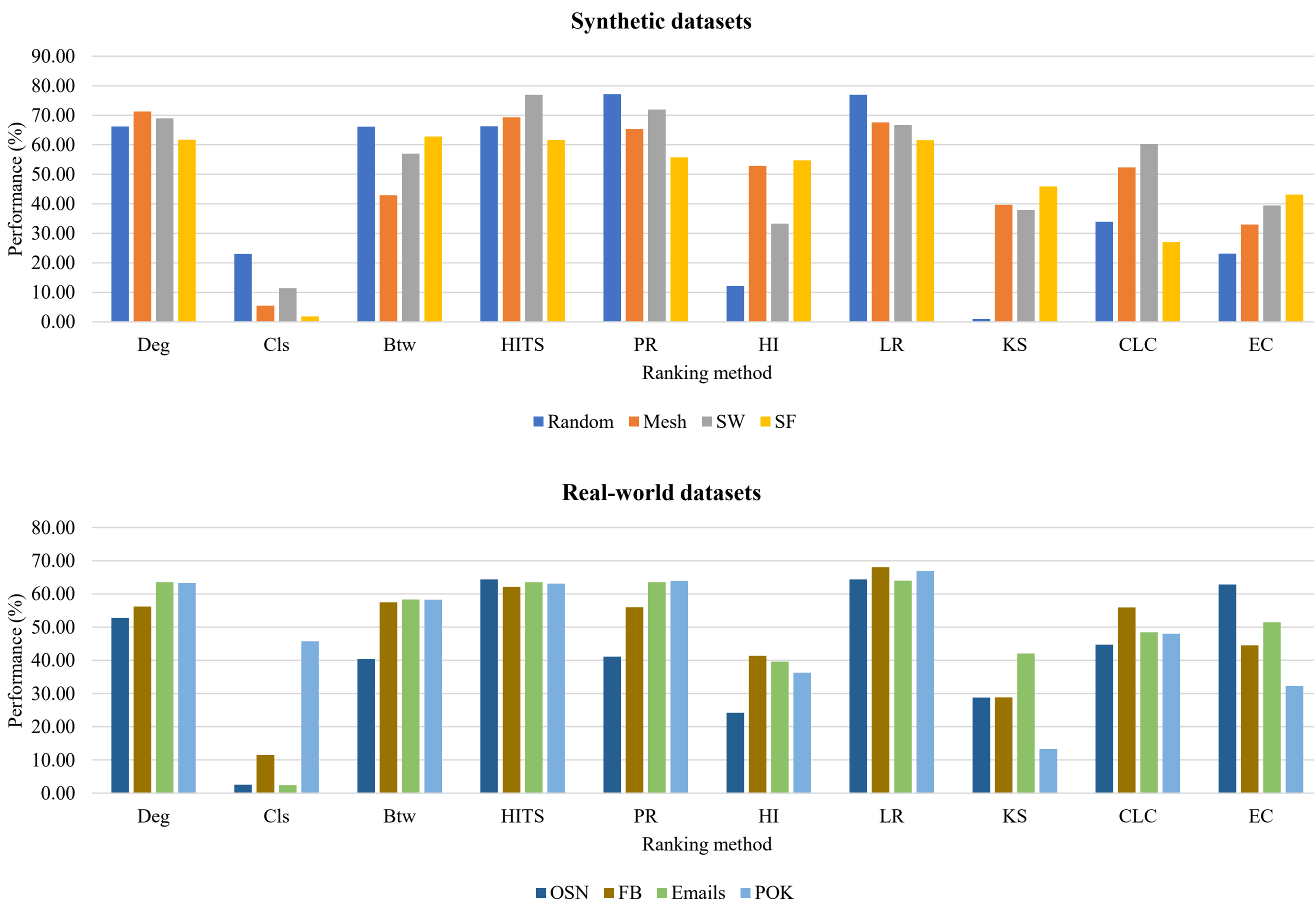

Figure 4. Performance (coverage 0-100%) of each ranking method on the 8 datasets using simultaneous competition-based benchmarking.

6 Alternative selection strategy for multiple spreaders

In order to present and validate our novel benchmarking methodology we have made use of existing state of the art ranking methods, as well as a selection strategy for multiple spreaders. After a review of the most recent advances in complex network analysis, we find that the method of simply selecting the top spreaders from the entire network is the most popular found throughout literature, in the case of multiple spreader selection[6,14–19]. As such, in terms of validating our benchmarking framework, we did not consider going into further details that were not the main goal of the paper.

Nevertheless, we find several alternatives for selecting multiple spreaders. Zhao *et al.* propose an innovative selection method using the Welsh-Powell graph colouring algorithm[20]. It is shown that their method can improve the performances of some well-known centralities, including degree, betweenness, closeness, and eigenvector centrality[21]. Nevertheless, there are other alternatives for multiple spreader selection as well, like recalculating the centralities of nodes after every step of node removal[22], the degree discount algorithm[23], the equal graph partitioning strategy[24].

We have implemented the graph colouring method of Zhao *et al.*[20] to present a brief comparative analysis of the impact it has in the overall benchmarking methodology and its results. Using the same three competitive simulation examples (Deg-LR, LR-Btw, Cls-HI) as in the *Results* section, Figure 4, from our manuscript, we highlight here in Figure 5 the differences between the two selection methods using visual examples. We refer to the simple method of selecting the top seeds from the entire network as the *naïve method*, and to the method of Zhao *et al.*[20] as the *graph colouring* method.

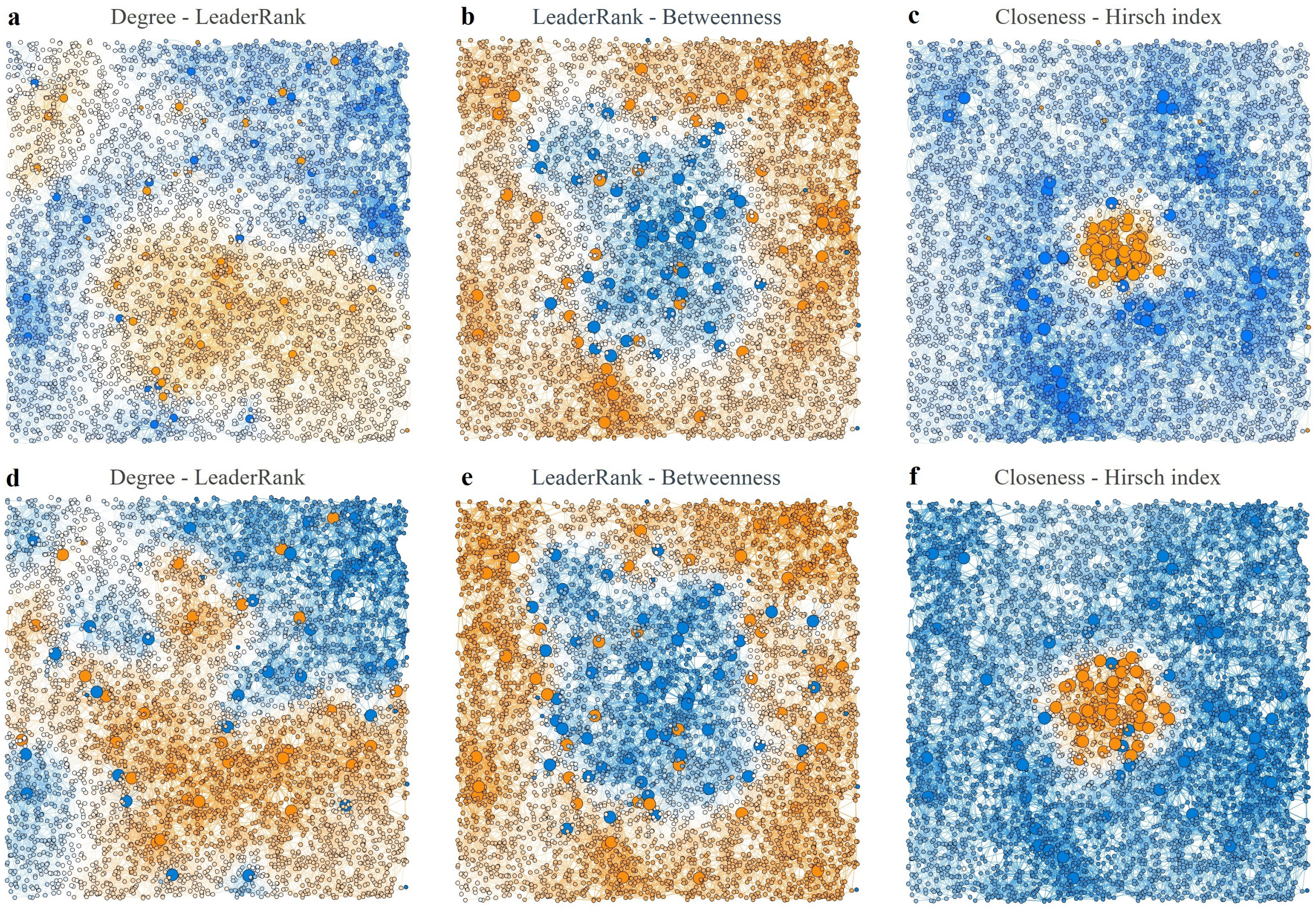

Figure 5. Comparison between the *naïve* (a-c) and *graph colouring* (d-f) methods using three competitive diffusion examples on the Mesh network (N=10,000 nodes). Larger nodes represent spreader nodes. The first centrality in the figure captions corresponds to orange opinion, and the second centrality to blue opinion.

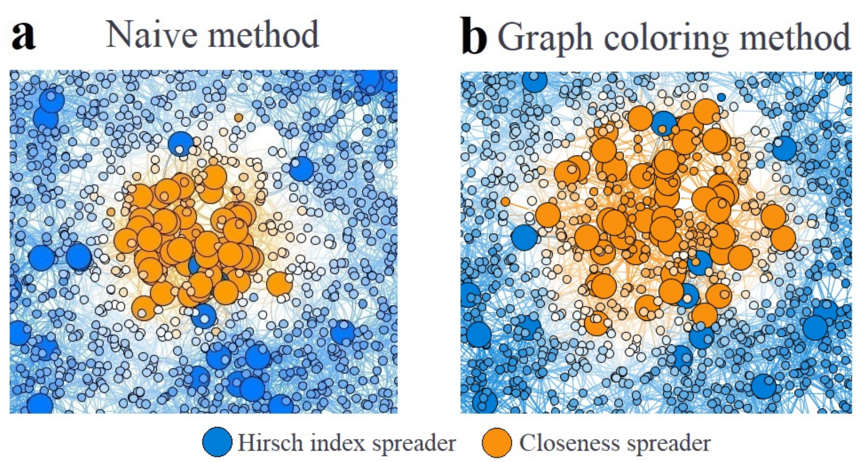

Figure 6. Difference in spreader spacing for closeness (orange) when switching from the *naïve* method (a) to the *graph colouring* method (b).

Analysing Figure 5 we notice a slightly different positioning of the spreader nodes (larger nodes) achieved with the *graph colouring* method. This is especially visible in Figures 5 c and f, where the spreaders selected by closeness are more spaced apart, yet the final distributions of the opinion (*i.e.*, centrality performances) are very similar when using either of the two selection methods. Additionally, in Figure 6 we provide a close-up of the spreader selection for Closeness (corresponding to Figures 5 c and f).

To sum up the results obtained for the two selection methods, in Table 3 we provide the numerical results for the three simulation scenarios. Overall, we measure variations in final performance (%) of roughly 1-3%.

Table 3. Comparison between the *naïve* and *graph colouring* methods in terms of selecting spreader nodes. Performance is expressed as percentage (%) for each node centrality in three competitive simulation scenarios.

	Ranking method	Naïve method	Graph colouring
1	Deg *versus*	56.70	53.66
	LR	43.30	46.34
2	LR *versus*	74.26	73.30
	Btw	25.74	26.70
3	Cls *versus*	5.24	6.82
	HI	94.76	93.18

References

1. Wang, X. F. & Chen, G. Complex networks: small-world, scale-free and beyond. *IEEE circuits systems magazine* **3**, 6–20 (2003).

2. Newman, M. E. The structure and function of complex networks. *SIAM review* **45**, 167–256 (2003).

3. Kleinberg, J. M. Authoritative sources in a hyperlinked environment. *J. ACM (JACM)* **46**, 604–632 (1999).

4. Page, L., Brin, S., Motwani, R. & Winograd, T. The pagerank citation ranking: Bringing order to the web. Tech. Rep., Stanford InfoLab (1999).

5. Hirsch, J. E. An index to quantify an individual's scientific research output. *Proc. Natl. academy Sci. United States Am.* 16569–16572 (2005).

6. Lü, L., Zhang, Y.-C., Yeung, C. H. & Zhou, T. Leaders in social networks, the delicious case. *PloS one* **6**, e21202 (2011).

7. Kitsak, M. *et al.* Identification of influential spreaders in complex networks. *Nat. physics* **6**, 888–893 (2010).

8. Zhao, X., Liu, F., Wang, J., Li, T. *et al.* Evaluating influential nodes in social networks by local centrality with a coefficient. *ISPRS Int. J. Geo-Information* **6**, 35 (2017).

9. Bonacich, P. Some unique properties of eigenvector centrality. *Soc. networks* **29**, 555–564 (2007).

10. Udrescu, M. & Topirceanu, A. Probabilistic modeling of tolerance-based social network interaction. In *Network Intelligence Conference (ENIC), 2016 Third European*, 48–54 (IEEE, 2016).

11. Topirceanu, A., Udrescu, M., Vladutiu, M. & Marculescu, R. Tolerance-based interaction: A new model targeting opinion formation and diffusion in social networks. *PeerJ Comput. Sci.* **2**, e42 (2016).

12. Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Phys. review letters* **86**, 3200 (2001).

13. Newman, M. E. Spread of epidemic disease on networks. *Phys. review E* **66**, 016128 (2002).

14. Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *nature* **406**, 378 (2000).

15. Cohen, R., Erez, K., Ben-Avraham, D. & Havlin, S. Breakdown of the internet under intentional attack. *Phys. review letters* **86**, 3682 (2001).

16. Pastor-Satorras, R. & Vespignani, A. Immunization of complex networks. *Phys. Rev. E* **65**, 036104 (2002).

17. Chen, D.-B., Gao, H., Lü, L. & Zhou, T. Identifying influential nodes in large-scale directed networks: the role of clustering. *PloS one* **8**, e77455 (2013).

18. Liu, J.-G., Lin, J.-H., Guo, Q. & Zhou, T. Locating influential nodes via dynamics-sensitive centrality. *Sci. reports* **6** (2016).

19. Li, Q., Zhou, T., Lü, L. & Chen, D. Identifying influential spreaders by weighted leaderrank. *Phys. A: Stat. Mech. its Appl.* **404**, 47–55 (2014).

20. Zhao, X.-Y., Huang, B., Tang, M., Zhang, H.-F. & Chen, D.-B. Identifying effective multiple spreaders by coloring complex networks. *EPL (Europhysics Lett.* **108**, 68005 (2015).

21. Lü, L. *et al.* Vital nodes identification in complex networks. *Phys. Reports* **650**, 1–63 (2016).

22. Holme, P., Kim, B. J., Yoon, C. N. & Han, S. K. Attack vulnerability of complex networks. *Phys. review E* **65**, 056109 (2002).

23. Chen, W., Wang, Y. & Yang, S. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 199–208 (ACM, 2009).

24. Chen, Y., Paul, G., Havlin, S., Liljeros, F. & Stanley, H. E. Finding a better immunization strategy. *Phys. review letters* **101**, 058701 (2008).