Influence Maximisation on Networks Predicted by Generative Models*

Valerio Restocchi[1], Soheil Eshghi[2], Setareh Magshoudi[2,3], Leandros Tassulias[2], Sebastian Stein[1]

[1] University of Southampton, Southampton, UK
[2] Yale University, New Haven, USA
[3] Technische Universität Berlin
[4] IBM T.J. Watson Research, Yorktown Heights, NY, USA
[5] Imperial College London, UK

1 Summary

Recently, the availability of large datasets on social networks has significantly facilitated the study of influence propagation, providing real-world benchmarks for validating theoretical models. Although this has become a prominent topic of research, little work analyses influence propagation under information constraints (i.e., partial and/or uncertain information). However, such constraints are common in real-world settings, and especially in coalition operations, where parts of an external group may be intentionally obscured by adversaries and where its structure must be inferred from noisy surveillance data.

We propose new approaches that exploit generative models of network topology to significantly improve influence diffusion in partially observable networks. We show that these algorithms perform similarly to or better than the state-of-the-art, showing promise for future research in this direction.

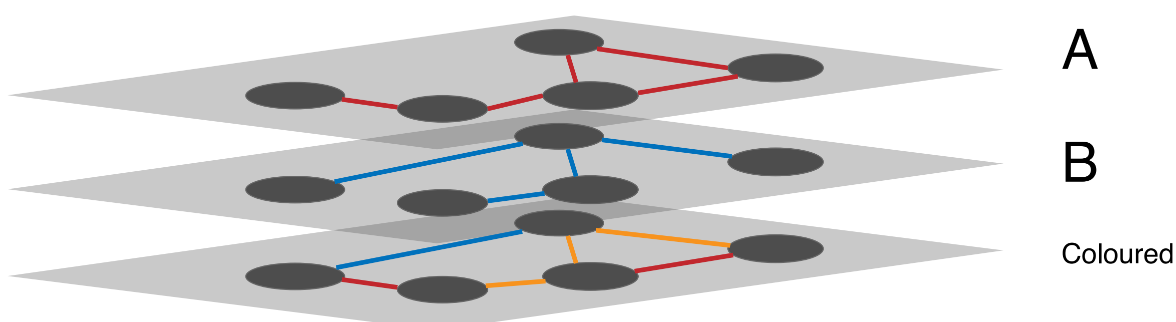


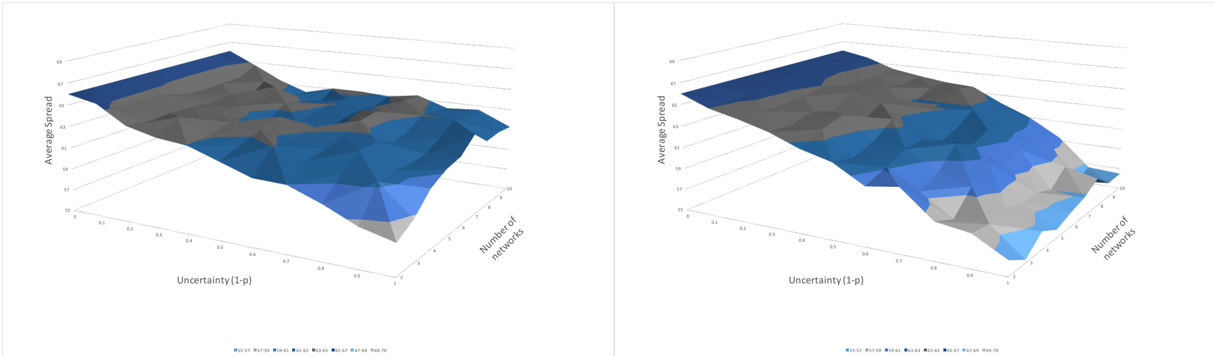
Fig. 1: Representation of a coloured network.

2 Influence maximisation on multiplex networks

The first scenario is that of partial observability. In this context, only a portion of a network is observable, and there is no information about the rest. However, we assume a generative model recreates N possible networks, by reconstructing the unobservable parts with accuracy $p \in [0,1]$. The accuracy p represents the probability that a given edge of the graph is part of the true, unobserved network.

We propose two heuristic algorithms for influence maximisation, namely Average Degree (AD) and Multiplex Weighted Degree (MWD), that exploit the idea of coloured networks [1, ?]. That is, given N=3 possible networks

*Part of this work is under review at AAAI.



(a) (b)

Fig. 2: Average spread for 5 nodes for AD (Figure a), and MWD (Figure b).

generated by a generative models, A, B, and C, we generate L=7 layers that represents networks in which links are either those exclusive to one of the three generated networks, present in both A and B, A and C, or B and C, or present in A, B, and C (see Figure 1). Both algorithms build the seed set by selecting the nodes with the highest average degree, considering all layers. Additionally, MWD assigns a weight w to each degree value, proportional to the number of networks in which the link is found. For example, MWD would assign a degree value of 4 to a node with 2 links in the layer that considers edges appearing in both A and B. This way, MWD exploits the intuition that the more networks an edge appears in, the higher its likelihood of being a real edge.

Figure 2 shows the results of our simulations, conducted on N synthetic scale-free networks with 1000 nodes. MWD outperforms AD only for low and intermediate levels of uncertainty. The fact that, for high levels of uncertainty MWD does not perform as well as with less noise, is likely because the algorithm assigns too high a weight to nodes whose edges appear in many networks only by chance.

3 Efficient Influence Maximisation Under Partial Network Visibility

In the second setting we investigate, we assume that a partial section of the network is observable, while the rest is one of a finite set of known structures, each with a given realisation probability. We show that solving the IM problem in this setting is NP-hard, and we provide analytical worst-case bounds for the performance of a novel computationally efficient approximation algorithm. In empirical experiments on real-world social networks, we demonstrate the efficiency of our algorithm and show that it outperforms state-of-the-art approaches that do not model the uncertainty. Finally, we show that computationally more efficient heuristics also perform well in some settings, albeit without the guarantees our algorithm provides.

References

1. F. Battiston, V. Nicosia, and V. Latora. Structural measures for multiplex networks. Physical Review E 89 (2014)
2. M. Kivelä, A. Arenas, M. Barthelemy, J. Gleeson, Y. Moreno, and M. Porter. Multilayer Networks. Journal of Complex Networks 2(3) (2014)