

MUTUAL REINFORCEMENT AT HIGHER-ORDER

Francesca Arrigo

Dept of Mathematics and Statistics
University of Strathclyde, UK
francesca.arrigo@strath.ac.uk

Desmond J. Higham

School of Mathematics
University of Edinburgh, UK
d.j.higham@ed.ac.uk

Francesco Tudisco

School of Mathematics
Gran Sasso Science Institute, Italy
francesco.tudisco@gssi.it

This 2–page abstract is based on *F. Arrigo, D. J. Higham and F. Tudisco, “A framework for second-order eigenvector centralities and clustering coefficients”, Proceedings of the Royal Society A, 476:20190724, 2020*

1 Introduction

Pairwise node-node interactions are the building block of networks and network mining algorithms. The adjacency matrix A can be used to represent these interactions. If there are n nodes, A is a $n \times n$ matrix with $A_{ij} = 1$ if nodes i and j are connected and $A_{ij} = 0$ otherwise. We assume here undirected edges so that A is symmetric.

In order to motivate our work, we consider here the eigenvector centrality. This is a vector u whose entries quantify the importance of the nodes in terms of the Perron eigenvector of the adjacency matrix A , precisely

$$\lambda u_i = \sum_{j:j \sim i} u_j = \sum_{j=1}^n A_{ij} u_j \quad \lambda \in \mathbb{R}, \lambda > 0, \quad u \in \mathbb{R}^n, u > 0$$

This centrality measure is *mutually reinforcing*, in the sense that the importance of node i is defined in terms of the importances of its neighbors. The same mutual reinforcement concept is at the core of all centrality measures based on eigen and singular vectors of graph matrices, as for example HITS and Google’s PageRank algorithms.

Mutually reinforcing algorithms are very useful and widely used. However, the majority of these algorithms consider only first-order neighborhoods. On the other hand, it is becoming apparent that many important network features arise from the interaction of larger groups of nodes [1, 4, 6, 7]. In particular, while information on higher-order interactions among nodes is *indirectly* used in many network science algorithms by considering traversals around the network, recent work has shown that there is a benefit in *directly* taking into account this information when designing graph algorithms.

In this talk we present a general tensor-based framework for incorporating second-order features (e.g. triangles) into mutually reinforcing network measures.

2 Nonlinear eigenvector framework

We propose and analyze a new constrained nonlinear eigenvector problem of the form

$$\lambda u = \alpha M u + (1 - \alpha) T_p(u) \quad \lambda \in \mathbb{R}, \lambda > 0, \quad u \in \mathbb{R}^n, u > 0 \quad (1)$$

where α is a coefficient in $(0, 1)$, M is a graph matrix (e.g. adjacency or PageRank matrix) and T_p is the mapping

$$T_p : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad x \mapsto T_p(x)_i = \sum_{j,k=1}^n T_{ijk} \left(\frac{|x_j|^p + |x_k|^p}{2} \right)^{1/p}$$

The coefficients T_{ijk} form a cubic tensor that takes into account second-order graph interactions. For example, T_{ijk} may be the binary triangle tensor, where $T_{ijk} = 1$ if ijk form a triangle in the network and $T_{ijk} = 0$ otherwise.

A vector u solution to (1) models a mutual reinforcement property u_i : each node i inherits additional importance from the nodes in its second-order neighborhood. For example, when T is the triangle tensor, u_i is large if i takes part in triangles that involve important nodes.

The map T_p is defined in terms of a parameter $p \in \mathbb{R}$ which allows us to tune the way nodes inherit the importances from their second-order neighbors. For example, if $p \rightarrow 0$ we obtain the geometric mean $\sqrt{|x_j x_k|}$ while for $p \rightarrow \infty$ we get $\max\{|x_j|, |x_k|\}$. The parameter α , instead, interpolates between traditional edge-based mutual reinforcement, for $\alpha = 1$, and a purely second-order model, obtained for $\alpha = 0$.

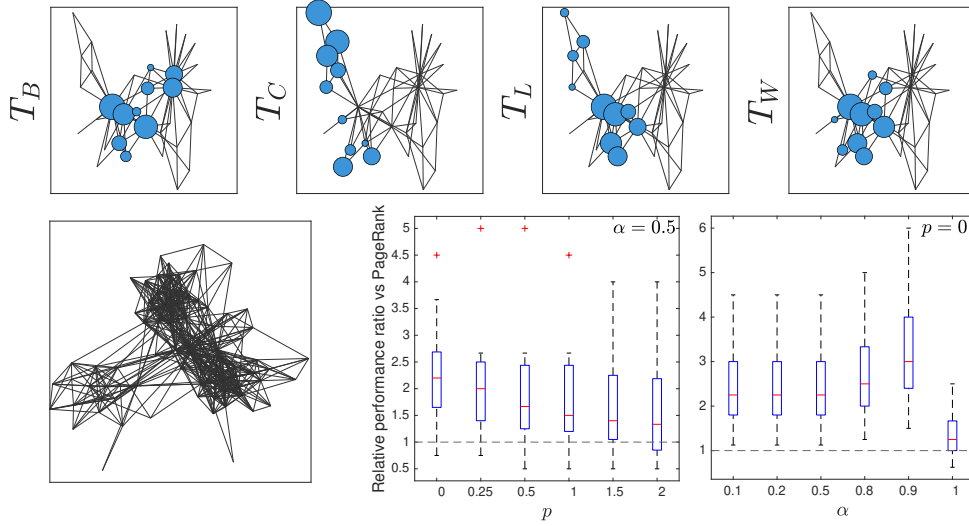


Figure 1: Top: 10 most central nodes in the Karate club network for different second-order eigenvector centrality models. Bottom: Link-prediction performance ratio between second- and first-order seeded PageRank algorithms.

Different choices of M and T give rise to different mutual reinforcing coefficients. For example, M can be the PageRank matrix and T the multilinear PageRank tensor [3]. If we define $T_{ijk} = 1/(d_i(d_i - 1))$ if ijk form a triangle and $T_{ijk} = 0$ otherwise, we obtain a spectral version of the Watts–Strogatz clustering coefficient [8]. Here d_i denotes the degree of node i . Letting $\omega_i = \sum_{j:j \sim i} (d_j - 1)$, the choice $T_{ijk} = 1/\omega_i$ if ijk form a triangle and $T_{ijk} = 0$ yields a mutual reinforcing version of the local closure coefficient [9].

3 Results

Using recent work on nonlinear Perron–Frobenius theory [2], we prove existence of a unique solution to (1), under mild assumptions on the topology of the graph. Moreover, we show that the solution to (1) can be computed efficiently using a *nonlinear power method*. As for the standard power method, the new method is guaranteed to converge for any positive starting point, the convergence rate is typically linear and the cost of each iteration is dominated by the cost of applying T_p . Hence the method scales to large and sparse networks that arise in many applications.

Figure 1 illustrates the computational results that will be presented in the talk. The first row shows the top 10 most central nodes in the Karate club network, obtained using the mutual reinforcing centrality in (1) for different choices of the tensor T . The second row of the figure compares link-prediction performances on a citation network between a second-order version of seeded PageRank obtained as a solution to (1) and the standard seeded PageRank algorithm [5]. The plot shows median and quartiles of the ratio of correctly predicted edges.

References

- [1] A. R. Benson, D. F. Gleich, and J. Leskovec. Higher-order organization of complex networks. *Science*, 353:163–166, 2016.
- [2] A. Gautier, F. Tudisco, and M. Hein. The Perron-Frobenius theorem for multihomogeneous mappings. *SIAM Journal on Matrix Analysis and Applications*, 40:1179–1205, 2019.
- [3] D. F. Gleich, L.-H. Lim, and Y. Yu. Multilinear PageRank. *SIAM J. Matrix Analysis and Applications*, 36:1507–1541, 2015.
- [4] A. P. Kartun-Giles and G. Bianconi. Beyond the clustering coefficient: A topological analysis of node neighbourhoods in complex networks. *Chaos, Solitons & Fractals: X*, 1:100004, 2019.
- [5] D. Liben-Nowell and J. Kleinberg. The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology*, 58(7):1019–1031, 2007.
- [6] C. E. Tsourakakis, J. Pachocki, and M. Mitzenmacher. Scalable motif-aware graph clustering. In *Proceedings of the International Conference on World Wide Web*, pages 1451–1460, 2017.
- [7] F. Tudisco, A. R. Benson, and K. Prokopcik. Nonlinear higher-order label spreading. *arXiv:2006.04762*, 2020.
- [8] D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393:440–442, 1998.
- [9] H. Yin, A. R. Benson, and J. Leskovec. The local closure coefficient: a new perspective on network clustering. In *Proceedings of the International Conference on Web Search and Data Mining*, pages 303–311. ACM, 2019.