On PageRank centrality for temporal networks

D. Aleja ♭ ♮, J. Flores♭, E. Primo♭ ♮ [1] and M. Romance♭ ♮

(♭) Department of Applied Mathematics, Science and Engineering of Materials, and Electronic Engineering, ESCET, Universidad Rey Juan Carlos, Móstoles (Madrid, Spain).

(♮) Laboratory of Computational Mathematics on Complex Networks and their Applications, Universidad Rey Juan Carlos, Móstoles (Madrid, Spain).

Abstract

The celebrated PageRank algorithm is at the core of the most popular commercial web page search engine and it is based in the heuristic idea that if we surf at random on a complex networks, the more frequent a node is visited, the more relevant it is [4]. More precisely, if we take a directed graph $\mathcal{G} = (V, E)$ of n nodes, we consider a *damping factor* $\lambda \in (0, 1)$ and two probability distribution vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{\mathbf{n}}$, ($\mathbf{v}$ is called a *personalization vector* and $\mathbf{u}$ is related to the dangling nodes of $\mathcal{G}$), then the *Google matrix*, $G = G(\lambda, \mathbf{u}, \mathbf{v})$, with $\lambda \in (0, 1)$, is the primitive and row-stochastic matrix defined as

$$G = \lambda(P_A + \mathbf{d}\mathbf{u}^T) + (1 - \lambda)\mathbf{e}\mathbf{v}^T \in \mathbb{R}^{n\times n}, \tag{1.1}$$

where P_A is the row-stochastic normalization of A, the adjacency matrix of $\mathcal{G}$, and $\mathbf{d}$ is the vector with entries 0 or 1 encoding the dangling nodes. The *Personalized PageRank vector* $\pi = \pi(\lambda, \mathbf{u}, \mathbf{v})$ (in the sequel *the PPR vector*) is the unique positive eigenvector of G^T associated to the eigenvalue 1 (see [4]).

In this talk we consider the Personalized PageRank centrality but for temporal networks, i.e. a family of graphs $\mathcal{G}(t) = (V, E(t))_{t\in I}$ with a fixed set of nodes on a time-scale I, either a discrete time-set or a (continuous) interval in the reals, and where all three A, $\mathbf{v}$ and $\mathbf{d}$ are time-dependent. In this temporal setting the associated time-dependent row stochastic Google matrix looks like

$$G(t) = \lambda(P_A(t) + \mathbf{d}(t)\mathbf{u}^T) + (1 - \lambda)\mathbf{e}\mathbf{v}(t)^T \in \mathbb{R}^{n\times n}.$$

Based on [2], a definition of Personalized PageRank vector is proposed that will take into account the interactions between nodes in past stages in coherence with the model introduced in [1]. Furthermore, an analytical *localization* result is presented in the talk that quantifies the influence of $\mathbf{v}(t)$ in the ranking of a each node. This new approach can be applied to both discrete and continuous temporal networks, showing that in the continuous setting the PPR vector can be estimated by the PPR vectors of the discrete samplings as it will be shown to be the limit of those.

References

[1] Ahmad, W., Porter, M. and Beguerisse-Díaz, M., Tie-Decay Networks in Continuous Time and Eigenvector-Based Centralities, *IEEE Trans. Networks Science and Eng.*, 8(2)1759—1771, 2021.

[2] Flores, J. and Romance, M., On eigenvector-like centralities for temporal networks: discrete vs. continuous time scales, *J. Comp. and Appl. Math.*, 330:1041–1051, 2018.

[3] García, E., Pedroche, F. and Romance, M., On the localization of the personalized PageRank of complex networks, *Lin. Alg. Appl.*, 439(3):640–652, 2013.

[4] Page, L. at al., The PageRank Citation Ranking: Bringing Order to the Web, *Technical Report* SIDL-WP-1999-0120, Stanford Digital Library Technologies Project, 1999.

[1] eva.primo@urjc.es