Fast Computation of Centrality Indices

Caterina Fenu*
Department of Computer Science, University of Pisa, caterina.fenu@for.unipi.it

James Baglama
Department of Mathematics, University of Rhode Island, jbaglama@math.uri.edu

David Martin
Department of Mathematical Sciences, Kent State University, dmarti49@kent.edu

Lothar Reichel
Department of Mathematical Sciences, Kent State University, reichel@math.kent.edu

Giuseppe Rodriguez
Department of Mathematics and Computer Science, University of Cagliari, rodriguez@unica.it

Abstract. One of the main issues in complex networks theory is to find the "most important" nodes. To this aim, one can use matrix functions applied to its adjacency matrix. After an introduction on the use of Gauss-type quadrature rules, we will discuss a new computational method to rank the nodes of both directed and undirected (unweighted) networks according to the values of these functions. The algorithm uses a low-rank approximation of the adjacency matrix, then Gauss quadrature is used to refine the computation.

Important Nodes

Graphs and Complex Networks are used to model interactions between various entities in real life applications, e.g. in computer science, sociology, economics, genetics, epidemiology. A graph is a pair of sets $G = (V, E)$, with $|V| = n$ and $|E| = m$. The elements of V are called nodes or vertices and those of E are known as edges or arcs. If the edges can be travelled in both directions the network is said to be undirected, directed otherwise. The adjacency matrix corresponding to an unweighted graph is the matrix $A \in \mathbb{R}^{n\times n}$ such that $A_{ij} = 1$ if there is an edge from node i to node j, and $A_{ij} = 0$ if node i and j are not adjacent. This kind of matrix is binary and, in general, nonsymmetric. One of the main issues in Complex Networks Theory is to find the "most important" nodes within a graph G. To this aim, various indices (or metrics) have been introduced to characterize the importance of a node in terms of connection with the rest of the network. The simplest and most classical ones are the *in-degree* and the *out-degree*, that is, the number of nodes that can reach one node or that can be reached from that node, respectively. These metrics do not give global information on the graph, since they only count the number of neighbors of each node. We will focus on indices that can be computed in terms of matrix functions applied to the adjacency matrix of the graph. We can define a class of indices starting from a matrix function

$$f(A) = \sum_{m=0}^{\infty} c_m A^m, \qquad c_m \geq 0.$$

Since $[A^m]_{ij}$ gives the number of paths of length m starting from the node i and ending at node j, $[f(A)]_{ij}$ is a weighted average of all the paths connecting i to j, and describes the ease of travelling between them. We refer to $[f(A)]_{ii}$ as the f-centrality of node i, and $[f(A)]_{ij}$ as the f-communicability between node i and node j.

In the literature, particular attention has been reserved to the exponential function. In [4], [5], the authors refer to $\left[e^A\right]_{ii}$ as the *subgraph centrality* of node i and to $\left[e^A\right]_{ij}$ as the *subgraph communicability* between node i and node j, in the case of an undirected graph. Recently, the notion of *hub centrality* and *authority centrality* has been introduced [3] in the case of a directed graph.

Benzi and Boito [2], following the techniques described by Golub and Meurant [8], employed quadrature formulas to find upper and lower bounds of bilinear forms of the kind $\mathbf{u}^T f(A)\mathbf{v}$ (with $\mathbf{u}$ and $\mathbf{v}$ unit vectors) in the case of a symmetric adjacency matrix.

If we assume that $[f(A)]_{ii}$ is a measure of the importance of node i, then we can identify the m most important nodes as the m nodes with the largest centrality. In order to do this using Gauss-type quadrature rules, we may apply this method with $\mathbf{u} = \mathbf{e}_i$, for $i = 1, \ldots, n$. Since a complex network is generally very large, this approach may be impractical.

Low-rank approximation

We will describe a new computational method to rank the nodes of both undirected and directed unweighted networks, according to the values of the above matrix functions. The idea is to reduce the cardinality of the set of candidate nodes in order to apply Gauss-type quadrature rules to a smaller number of nodes. The first part of the resulting algorithm, called *hybrid method*, is based on a low-rank approximation of the adjacency matrix. If the network is undirected a partial eigenvalue decomposition is used [6], while if the network is directed we make use of a partial singular value decomposition [1]. We will compare the hybrid algorithm to other computational approaches, on test networks coming from real applications, e.g. in software engineering, bibliometrics and social networks. We will also present a block algorithm to compute the entries of the matrix exponential for both symmetric and nonsymmetric matrices, which is particularly efficient on computers with a hierarchical memory structure. This algorithm is based on block Gauss and anti-Gauss quadrature rules. In the case of a nonsymmetric matrix the block approach is necessary to avoid breakdowns during the computation [7], [9].

References

[1] J. Baglama, C. Fenu, L. Reichel, G. Rodriguez, *Analysis of directed networks via partial singular value decomposition and Gauss quadrature*, Linear Algebra Appl. 456 (2014), 93–121.

[2] M. Benzi, P. Boito, *Quadrature rule-based bounds for functions of adjacency matrices*, Linear Algebra Appl. 433 (2010), 637–652.

[3] M. Benzi, E. Estrada, C. Klymko, *Ranking hubs and authorities using matrix functions*, Linear Algebra Appl. 438 (2013), 2447–2474.

[4] E. Estrada, N. Hatano, *Communicability in complex networks*, Phys. Rev. E 77 (2008), 036111.

[5] E. Estrada, J. A. Rodríguez-Velázquez, *Subgraph centrality in complex networks*, Phys. Rev. E 71 (2005), 056103.

[6] C. Fenu, D. Martin, L. Reichel, G. Rodriguez, *Network analysis via partial spectral factorization and Gauss quadrature*, SIAM J. Sci. Comput. 35 (2013), A2046–A2068.

[7] C. Fenu, D. Martin, L. Reichel, G. Rodriguez, *Block Gauss and anti-Gauss quadrature with application to networks*, SIAM Journal on Matrix Analysis and Applications 34(4) (2013), 1655–1684.

[8] G. H. Golub, G. Meurant, *Matrices, Moments and Quadrature with Applications.* Princeton University Press, Princeton, 2010.

[9] D. R. Martin, *Quadrature approximation of matrix functions, with applications*, PhD Thesis, 2012.