HORIZON 2020

Models and Algorithms for Graph centrality grounded on Nonlinear Eigenvalues Techniques

Reporting

Project Information

MAGNET

Grant agreement ID: 744014

Project website

DOI
10.3030/744014

Project closed

EC signature date
20 March 2017

Start date
1 July 2017

End date
30 June 2019

Funded under
EXCELLENT SCIENCE - Marie Skłodowska-Curie Actions

Total cost
€ 183 454,80

EU contribution
€ 183 454,80

Coordinated by
UNIVERSITY OF STRATHCLYDE
United Kingdom

Periodic Reporting for period 1 - MAGNET (Models and Algorithms for Graph centrality grounded on Nonlinear Eigenvalues Techniques)

Reporting period: 2017-07-01 to 2019-06-30

Summary of the context and overall objectives of the project

"Networks are fundamental tools for the analysis of complex systems. In fact, we often state questions and develop analysis in terms of nodes and edges. For example, we can describe a social system by modeling individuals as nodes and social interactions as edges between pair of nodes. Similarly, we can model a biochemical reaction by assigning a node to each protein and edges between them to model physical contacts of high specificity, so-called protein-protein interactions. Due to the broad scope of this modeling paradigm,
the analysis of systems or datasets as networks has enjoyed tremendous success over the last decade.

This project has focused on the task known as network centrality where one wants to quantify the importance of the nodes of a network by exploiting only its topological structure, i.e.\ the nodes and the edges between them, without assuming any other external information about the data.

Being able to identify important nodes is a very relevant issue which is useful in numerous contexts. For example, it can be used to identify the most important individuals in a communication system and thus maximize the effect of commercial campaigns, to locate infected nodes in a population with diseases and thus to predict the spreading of diseases, to improve the ordering of the results of search engines, to identify common patterns in digital datasets to improve text-to-speech computer systems, or to discover and predict anomalous financial quantity behaviors.

From a mathematical point of view, network centrality requires to address two main issues: to design the model that defines the importance of the nodes and to develop algorithms that allow to efficiently compute such importances on large network datasets.

One of the oldest and most successful centrality models, known as ""eigenvector centrality"", is based on the following LINEAR mutual reinforcing argument: the importance of a node is proportional to the sum of the importances of the nodes it is connected to. For example, this is the model used by the Google search engine to decide the ordering of the pages result of a search query. Together with its well understood mathematical model, the success of eigenvector centrality is due to a simple algorithm that can be used for its computation: the so-called ""power method"".

However, because of advances in technology and the increased digitisation of
human behaviour, network data is growing rapidly in terms of size and variety and the linearity of the eigenvector centrality model is a big limitation that prevents to fully exploit the complexity of modern network data.

This project had two major overall objectives:

1) Introduce new nonlinear eigenvector centrality models able to capture network structure details and complex higher-order node interactions that are overlooked by simpler linear models.
2) Design bespoke nonlinear versions of the power method in order to efficiently compute the newly introduced nonlinear centrality scores for large real-world complex networks."

Work performed from the beginning of the project to the end of the period covered by the report and main results achieved so far

We started by identifying two major settings where linear node importance models are not satisfactory. The first one is the case of modern multilayer networks, the second one is the evaluation of the coreness score of nodes and the corresponding detection of core-periphery structures.

We then started analyzing the plethora of recent work where models for node centrality on multilayer networks have been proposed. We found out that all the available models propose different ways preprocessing and simplifying the data by transforming the input multilayer structure into a standard mono-layer network and then apply to the new simplified dataset a standard eigenvector centrality techniques. We then developed a new and mathematically rigorous nonlinear eigenvector centrality model that allows to quantify the importance of the nodes in a multilayer network using the whole multilayer structure as it is, without any initial transformation. Actually, our model exploits all the information in the multiple layers and allows to quantify the relevance of different layers as well. Finally, we developed a nonlinear power method for this new centrality model and we proved convergence and quality guarantees for this method.

In a second phase, we studied the case of coreness importance. A thorough analysis of the available literature showed that most of the current approaches are based on the optimization of different types of core-quality kernel functions. However, they all require linear or quadratic kernels which result in simplistic and imprecise models. Thus, we have developed a model based on a family of nonlinear kernels that we have proved to be nearly optimal. Moreover, we have developed a bespoke version of the power method for this problem which we have proved, as for the previous case, to be globally convergent.

Progress beyond the state of the art and expected potential impact (including the socio-economic impact and the wider societal implications of the project so far)

The project has allowed the fellow to receive a unique academic training and he has now secured a tenure-track academic position in the competitive EU market.
The project has allowed to develop the notion of nonlinear eigenvector centrality and popularize it throughout both the network science and the numerical linear algebra communities. A major advancement was the realization that nonlinear eigenvectors and nonlinear Perron-Frobenius theory can be used in the context of network quality function optimization, which allows to optimally solve a family of nonconvex network-related optimization problems that are otherwise hardly tractable.

Another exciting development was the relation between the new centrality model and a class of nonlinear eigenvector problems (also called nonlinear eigenvalue problems with eigenvector nonlinearities) which we were able to solve with high precision thanks to specialized nonlinear power iterations.

The breakthrough advancement made possible by this nonlinear eigenvector approach to optimization and network science problems will have important long term impact in the fields of numerical analysis and mathematics of data science.

Core train stations in the city of London

Last update: 6 November 2019

Permalink: https://cordis.europa.eu/project/id/744014/reporting