Artificial Intelligence 231 (2016) 39–63

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

Efficient algorithms for game-theoretic betweenness centrality

Piotr L. Szczepański [a,*], Tomasz P. Michalak [b,c], Talal Rahwan [d]

[a] Institute of Informatics, Warsaw University of Technology, Pl. Polit. 1, 00-661 Warsaw, Poland
[b] Department of Computer Science, University of Oxford, OX1 3QD, UK
[c] Institute of Informatics, University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland
[d] Masdar Institute of Science and Technology, Abu Dhabi 54224, United Arab Emirates

A R T I C L E   I N F O

Article history:
Received 11 October 2014
Received in revised form 23 October 2015
Accepted 2 November 2015
Available online 5 November 2015

Keywords:
Betweenness centrality
Shapley value
Semivalue

A B S T R A C T

Betweenness centrality measures the ability of different nodes to control the flow of information in a network. In this article, we extend the standard definition of betweenness centrality using *Semivalues*—a family of solution concepts from cooperative game theory that includes, among others, the Shapley value and the Banzhaf power index. Any Semivalue-based betweenness centrality measure (such as, for example, the Shapley value-based betweenness centrality measure) has the advantage of evaluating the importance of individual nodes by considering the roles they each play in different groups of nodes. Our key result is the development of a general polynomial-time algorithm to compute the Semivalue-based betweenness centrality measure, and an even faster algorithm to compute the Shapley value-based betweenness centrality measure, both for weighted and unweighted networks. Interestingly, for the unweighted case, our algorithm for computing the Shapley value-based centrality has the same complexity as the best known algorithm for computing the standard betweenness centrality due to Brandes [15]. We empirically evaluate our measures in a simulated scenario where nodes fail simultaneously. We show that, compared to the standard measure, the ranking obtained by our measures reflects more accurately the influence that different nodes have on the functionality of the network.

1. Introduction

Betweenness centrality is a measure introduced independently by Anthonisse [3] and by Freeman [31], based on which the importance (or "centrality") of a node, v, is assumed to increase with the percentage of shortest paths (from all nodes to all other nodes) that pass through v. Everett and Borgatti [29] extended this measure to *groups* of nodes, whereby the importance of any such group increases with the percentage of shortest paths that pass through that group; this is called the *group Betweenness centrality measure*.

Several applications of betweenness centrality have been considered in the literature. Among others, Puzis et al. [55] studied how group betweenness centrality can guide the process of deploying intrusion-detection devices in a complex network. Dolev et al. [24] developed an extension of the standard measure, called routing betweenness centrality, to identify key routers in a communication network. Girvan and Newman [34] proposed a community-detection algorithm based on edge betweenness centrality. Furthermore, according to the European Central Bank [28], betweenness centrality is arguably the most suitable, simple measure of systemic importance of institutions in the financial system.

* Corresponding author. Tel.: +48 692144559.
E-mail addresses: szczep.pl@gmail.com (P.L. Szczepański), Tomasz.Michalak@cs.ox.ac.uk (T.P. Michalak), trahwan@masdar.ac.ae (T. Rahwan).

http://dx.doi.org/10.1016/j.artint.2015.11.001
0004-3702/

Table 1
Summary of the computational results obtained in this article. The complexity results, except for group betweenness centrality, concern computing the entire cardinal ranking of nodes. The complexity of group betweenness centrality concerns computing the centrality of a single set of nodes.

Centrality	Unweighted graphs	Weighted graphs
Standard betweenness	$O(\|V\|\|E\|)$, Brandes [15]	$O(\|V\|\|E\| + \|V\|^2 \log(\|V\|))$, Brandes [15]
Betweenness of a particular group	$O(\|V\|\|E\|)$, Brandes [16]	$O(\|V\|\|E\| + \|V\|^2 \log(\|V\|))$, Brandes [16]
Shapley value-based betweenness	$O(\|V\|\|E\|)$, **this article**	$O(\|V\|^2\|E\|) + \|V\|^2 \log \|V\|$, **this article**
Semivalue-based betweenness	$O(\|V\|^2\|E\|)$, **this article**	$O(\|V\|^3\|E\|) + \|V\|^3 \log \|V\|$, **this article**

In this article, we propose a generalization of betweenness centrality based on solution concepts from cooperative game theory. Our work falls under a wider body of research in which networks are analyzed using the combinatorial structure of coalitional games (see the next section for more details). The key advantage of this approach is that nodes are ranked not only based on their individual roles in the network but also based on their contribution to the role played by all possible subsets (or groups) of nodes. This is particularly relevant to settings in which the role played by a group of nodes is not simply the sum of the individual roles played by the group members. Consider, for instance, an epidemiology application, where the aim is to contain the spread of a disease [25]. If we ask the question of whether the vaccination of any individual node is sufficient to stop the spread of the disease then the answer is most probably negative. A much more likely way to contain the disease is to simultaneously vaccinate a group of carefully-selected nodes. In such a scenario, to quantify the importance of a node, one needs to consider the potential gain of vaccinating that node as part of a group, rather than just considering the potential gain of vaccinating it alone.

Such an analysis of groups of nodes in the network directly corresponds to coalitional game theory, where the performance of players is studied in "coalitions" (i.e., subsets of players). By imposing the combinatorial structure of a coalitional game over a network, it becomes possible to analyze the performance of nodes using a plethora of game-theoretic solution concepts, developed over decades to analyze the performance of players in a cooperative game. Note that a cooperative game typically places no assumptions or restrictions on how the groups are evaluated. Thus, when using such a game to model the roles played by nodes in a network, the group-evaluation function can be tailored to best fit the centrality measure at hand. For instance, a group of nodes can be evaluated based on the average degree therein, or based on its closeness to other nodes, etc. (more on this in the next section). As such, several game-theoretic network centralities have been proposed in the literature to date, each based on a different group-evaluation function [62,44,48].

Compared to the standard centrality measures, a potential downside of *game-theoretic centrality measures* is that they are based on solution concepts that are themselves hard to compute. For instance, given a coalitional game defined over a network of $|V|$ nodes, a straight-forward computation of the *Shapley value*—a fundamental game-theoretic solution concept—requires considering all possible $2^{|V|}$ coalitions (i.e., groups) of nodes. This is clearly prohibitive for networks with hundreds, or even tens, of nodes. Indeed, certain game-theoretic centrality measures have been proven impossible to compute in time polynomial in the size of the network [49]. On the other hand, some positive computational results have also been found. In particular, Michalak et al. [48] analyzed various Shapley value-based extensions of degree and closeness centrality and showed that it is possible to compute those in polynomial time. Nevertheless, these positive results were only for certain game-theoretic extensions of degree centralities, which are computationally less challenging than betweenness centrality. In fact, no game-theoretic extension of betweenness centrality has been developed to date.

Given this, our contributions in this article can be summarized as follows:

- We propose a game-theoretic extension of Betweenness centrality, based on a family of solution concepts from cooperative game theory known as *Semivalues* [50]. This family includes two of the most fundamental solution concepts in the literature, namely the *Shapley value* [56] and the *Banzhaf Index* [9]. To put it in the context of measuring the systemic importance of financial institutions, the standard betweenness centrality measure, suggested by the European Central Bank [28] as a simple index of systemic importance, assumes that any crisis is initiated by a single insolvent institution. On the other hand, our Semivalue-based betweenness centrality measure takes into consideration the fact that a crisis in the financial network can be initiated by a group of insolvent institutions, with varying probability depending on the size of the group.
- Our main technical contribution is to propose polynomial time algorithms to compute the Semivalue-based betweenness centrality measure, and an even faster algorithm to compute the Shapley value-based betweenness centrality measure, both for weighted and unweighted networks. Interestingly, as shown in Table 1, for the unweighted case our algorithm for computing the Shapley value-based centrality has the same complexity as the best known algorithm for computing the standard betweenness centrality, due to Brandes [15]. In particular, both algorithms run in $O(|V||E|)$ time, where V is the set of nodes and E is the set of edges in the network.
- Finally, we compare the performance of the new measures with that of the standard betweenness-centrality measure in a scenario of simultaneous node failures (in the spirit of Holme et al. [39]). To this end, we quantify the functionality of the resulting network based on four well-known measures, namely the *Average Inverse Geodesic* [39], the *Average*