Game Theoretic Network Centrality: Exact Formulas and Efficient Algorithms

(Extended Abstract)

Karthik V Aadithya
The University of California, Berkeley
kv.aadithya@gmail.com

Balaraman Ravindran
Indian Institute of Technology Madras
ravi@cse.iitm.ac.in

ABSTRACT

The concept of centrality plays an important role in network analysis. Game theoretic centrality measures have been recently proposed, which are based on computing the Shapley Value (SV) of each node (agent) in a suitably constructed co-operative network game (for example see [1]). However, the naive method of exact computation of SVs takes exponential time in the number of nodes. In this paper, we develop analytical formulas for computing SVs of nodes for various kinds of centrality-related co-operative games played on both weighted and unweighted networks. These formulas not only provide an efficient and error-free way of computing node centralities, but their surprisingly simple closed form expressions also offer intuition into why certain nodes are relatively more important to a network.

Categories and Subject Descriptors: H.0 [Information Systems]: General; H.5.3 [Group and Organization Interfaces]: Evaluation/methodology

General Terms: Algorithms, Economics, Game Theory

Keywords: Game Theory, Co-operative Games, Shapley Value, Social Networks

1. INTRODUCTION

The question of which nodes and edges are most important to a network arises naturally in many different contexts - such as determining the most influential people in a social network or the most important highways in a road network or the most critical functional entities in a protein network. As a result, the concept of centrality has been studied extensively in network analysis [2]. Here we explore notions of centrality based on game theoretic ideas.

Traditional centrality measures have usually worked by assigning a score to each node of a network, which in some way corresponds to the importance of that node for the application at hand. For instance, if the application is to design a communications network robust to targeted attacks, a traditional centrality measure might work by analysing the consequences of failure of each node. The more adverse the consequences of failure, the higher the node centrality. Different traditional measures of centrality have been evolved for different applications. These include degree centrality, betweenness centrality, closeness centrality, eigenvalue centrality etc [2].

Cite as: Game Theoretic Network Centrality: Exact Formulas and Efficient Algorithms (Extended Abstract), Karthik V Aadithya and Balaraman Ravindran, *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. 1459-1460

However, in many network applications, it is not sufficient to merely view nodes as stand-alone entities. One would ideally like to understand the importance of each node in terms of its utility when combined with other nodes [3]. For instance, in the above communications network, one would ideally like to assign a centrality score to a node n based on the consequences of failure of every arbitrary combination of nodes containing n, rather than just failure of the single node n.

Game theoretic centrality has been proposed as a framework that embraces the above philosophy [1]. Given the network to be analysed, the main idea here is to define a co-operative game involving the agents (nodes) of the network. This involves specifying the "coalition value" of every arbitrary combination of agents - which can depend in a highly non-trivial way on the underlying network structure. The final step is to determine the Shapley Value (SV) of each agent in this game, which is the required centrality measure.

The chief difficulty in adopting such a game theoretic centrality measure is that the naive algorithm for computing SVs in a coalitional game is of exponential complexity in the number of agents. For this reason, papers like [1] adopt a Monte-Carlo simulation based approximate method of evaluating SVs. However, the accuracy of such an approach is open to doubt because the simulator would only be able to explore a vanishingly small section of the sample space as the network becomes large.

Our key contribution is to show that it is possible to efficiently and accurately compute SVs in the context of many co-operative games of practical interest defined on large networks. Indeed, we develop exact closed-form formulas to characterise the SVs of agents in these games, which can be turned into linear and polynomial time algorithms to efficiently compute node centralities.

2. PRELIMINARIES

We assume that the reader is familiar with common notions of graph theory including weighted and unweighted graphs, vertex degrees, neighbouring nodes and shortest paths. We do not define these concepts here but suggest the reference [4]. All weighted graphs considered in this paper are positive weighted. Further, all graphs are assumed to be undirected. Also, we use $deg(n_i, G)$ to denote the degree of node $n_i \in V(G)$ in graph G. Lastly, the terms "graph" and "network" are used interchangeably in this paper, as are the terms "node" and "vertex".

We also assume that the reader is familiar with notions of co-operative game theory, including the definition of coalitional games and the Shapley Value. We do not define these concepts here but suggest the reference [5].

We now set the notation for a general coalitional game played on a graph. Given a graph $G(V, E)$ with vertex set V and edge set E, we define a coalitional game $g(V(G), v)$ with set of agents $V(G)$ and characteristic function v. That is, the agents (players) of the coalitional game are the vertices of the graph G. The characteristic function $v : 2^{V(G)} \to \mathbb{R}$ is a map from the powerset of $V(G)$ to the set of real numbers $\mathbb{R}$. The map v can be any function that depends on the graph G as long as it satisfies the condition $v(\emptyset) = 0$. We use the phrase "value of coalition s" to informally refer to $v(s)$ where $s \in 2^{V(G)}$.

Following the notation above, we now proceed to define and solve specific coalitional games played on weighted and unweighted graphs.

3. GAME 1

Given an unweighted graph G. We define $fringe(s, G)$ of a coalition $s \subseteq V(G)$ as the set $\{n_j \in V(G) : n_j \in s\,(or)\,(n_j, n_i) \in E(G)\ for\ some\ n_i \in s\}$. We observe that the fringe captures a "degree-like" intuition since it includes all nodes reachable from the coalition s in not more than one step.

We now define the coalitional game $g_1(V(G), v)$ by the characteristic function $v : 2^{V(G)} \to \mathbb{R}$ given by

$$v(\emptyset) = 0$$
$$v(s) = size(fringe(s, G))\ \forall s \neq \emptyset$$

This coalitional game has been extensively discussed in [1], where the authors motivate the game by arguing that the SVs of nodes in this game constitute a centrality metric that is superior to degree centrality for some applications. We shall now present an exact formula for computing these SVs (rather than using Monte-Carlo simulation as was done in [1]).

To evaluate the SV of node $n_i \in V(G)$, consider a permutation p of nodes of G, drawn uniformly at random from the set of all possible permutations. Let the set of nodes occurring before node n_i in p be denoted $S_p(n_i, G)$.

Now we ask the question - what is the necessary and sufficient condition for node n_i to marginally contribute node $n_j \in fringe(\{n_i\}, G)$ to $fringe(S_p(n_i, G), G)$? Clearly this happens if and only if neither n_j nor any of its neighbours are already present in $S_p(n_i, G)$. Given that p is chosen uniformly at random, the probability of this event occurring is given by $\frac{1}{1+deg(n_j,G)}$.

Let C_{n_i,n_j} denote the Bernoulli random variable that node n_i marginally contributes node $n_j \in fringe(\{n_i\}, G)$ to the value of coalition $S_p(n_i, G)$. Thus we have

$$E(C_{n_i,n_j}) = \frac{1}{1 + deg(n_j, G)}$$

But the Shapley Value $SV(n_i)$ is the expected total marginal contribution of node n_i given by

$$SV(n_i) = \sum_{n_j \in fringe(\{n_i\},G)} E(C_{n_i,n_j})$$

Thus we have

$$SV(n_i) = \sum_{n_j \in fringe(\{n_i\},G)} \frac{1}{1 + deg(n_j, G)}$$

which gives the required analytical expression for computing SV of each node.

Algorithm 1: Computing SVs for Game 1

Input: Undirected unweighted graph G
Output: SVs of all nodes in G wrt Game 1

foreach *node* $v \in V(G)$ **do**
  ShapleyValue[v] = $\frac{1}{1+deg(v,G)}$;
  foreach *neighbour* u *of* v **do**
    ShapleyValue[v] += $\frac{1}{1+deg(u,G)}$;
  end
end

Table 1: Other solved games

Graph	Characteristic function	Complexity
UW	$v(s)$ is the number of nodes that have atleast k neighbours belonging to s	$V + E$
UW	$v(s)$ is the number of nodes at-most k_{cutoff} degrees of separation away from some node in s	$VE + V^2 logV$
W	$v(s)$ is the number of nodes whose distance from some node in s is less than d_{cutoff}	$VE + V^2 logV$
W	$v(s)$ is an expression of the form $\sum_{n_i \in V(G)} f(\min_{n_j \in s} d(n_i, n_j))$ where $d(n_i, n_j)$ is the distance between nodes n_i and n_j and f is an arbitrary decreasing function of its argument.	$VE + V^2 logV$

$UW \to Unweighted;\ W \to Weighted$

Algorithm 1 computes the exact SVs of all nodes by implementing the above equation in $O(V + E)$ time. By contrast, Monte-Carlo simulation requires $O(V + E)$ computations for each iteration [1].

The above formula gives the intuition that a node will have high centrality not only when its degree is high, but also whenever its degree tends to be higher in comparison to the degree of its neighbouring nodes.

4. OTHER GAMES

The approach described in the previous section is not limited to game g_1 alone. In fact, it has also enabled us to derive analytical SV formulas for the more general scenarios outlined in the above table. The actual proofs and algorithms for these additional games are truly marvelous demonstrations. However, they require a significant amount of extra discussion and notation which this extended abstract is too small to contain.

5. REFERENCES

[1] N. Rama Suri and Y. Narahari. Determining the top-k nodes in social networks using the Shapley Value. In *AAMAS 2008*, 2008.
[2] Dirk Koschutzki, Katharina A. Lehmann, Leon Peeters, Stefan Richter, Dagmar Tenfelde-Podehl, and Oliver Zlotowski. *Centrality Indices*, pages 16–61. 2005.
[3] David Kempe, Jon Kleinberg, and Eva Tardos. Influential nodes in a diffusion model for social networks. In *ICALP*, 2005.
[4] JA Bondy and USR Murty. *Graph theory and its applications*. Macmillan, 1976.
[5] R. Gibbons. *A primer in game theory*. Harvester Wheatsheaf New York, 1994.