Centrality Analysis of the United States Network Graph

Natarajan Meghanathan
Jackson State University, USA

INTRODUCTION

Network Science is one of the emerging fields of Data Science to analyze real-world networks from a graph theory point of view. Several real-world networks have been successfully modeled as undirected and directed graphs to study the intrinsic structural properties of the networks as well as the topological importance of nodes in these networks. The real-world networks that have been subjected to complex network analysis typically fall under one of these categories: social networks (Ghali et. al., 2012), transportation networks (Cheung & Gunes, 2012), biological networks (Ma & Gao, 2012), citation networks (Zhao & Strotmann, 2015), co-authorship networks (Ding, 2011) and etc. One category of real-world networks for which sufficient attention has not yet been given are the regional networks featuring the states within a country. In this chapter, we present a comprehensive analysis of a network graph of the states within a country with respect to the four commonly used centrality metrics in complex network analysis (Newman, 2010): degree, eigenvector, betweenness and closeness centralities.

We opine the chapter to serve as a model for anyone interested in analyzing a connected graph of the states within a country from a Network Science perspective. The approaches presented in this chapter could be useful to determine the states (and their cities) that are the most central and/or influential within a country. For example, the ranking of the vertices based on the shortest path centrality metrics (closeness and betweenness) could be useful to choose the states (and their cities) that could serve as hubs for transportation networks (like road and airline networks). We could identify the states that are most the central states as well as identify the states that could form a connected backbone and geographically well-connected to the rest of the states within a country and use this information to design the road/rail transportation networks. The degree centrality and eigenvector centrality metrics as well as the network-level metrics like minimum connected dominating set and maximal clique size could be useful to identify fewer number of venues (with several adjacent states to draw people) for political campaigns/meetings that would cover the entire country.

We choose the United States (US) as the country for analysis and build a connected network graph of the contiguous states (48 states and the District of Columbia, DC) of the US: each state and DC is a node (vertex) and there exists a link (edge) between two vertices if the two corresponding states/DC share a common border. Though some prior studies have been conducted on transportation networks (Cheung & Gunes, 2012) and food flow networks (Lin et. al., 2014) in the United States, to the best of our knowledge, there has been no prior study of centrality analysis on the graph of the contiguous US states solely based on their geographical locations. Table 1 lists the contiguous states and DC in alphabetical order, their two character codes and the IDs used to refer to them in the chapter. The rest of the chapter is organized as follows: Section 2 introduces the centrality metrics and presents the results of the analysis on the states graph for each of them. Section 3 discusses related work. Section 4 concludes the chapter by summarizing the results

DOI: 10.4018/978-1-5225-2255-3.ch152

Table 1. List of contiguous states (including DC) of the US in alphabetical order

ID	State/District Name	Code	ID	State Name	Code	ID	State Name	Code
1	Alabama	AL	18	Maine	ME	34	Ohio	OH
2	Arizona	AZ	19	Maryland	MD	35	Oklahoma	OK
3	Arkansas	AR	20	Massachusetts	MA	36	Oregon	OR
4	California	CA	21	Michigan	MI	37	Pennsylvania	PA
5	Colorado	CO	22	Minnesota	MN	38	Rhode Island	RI
6	Connecticut	CT	23	Mississippi	MS	39	South Carolina	SC
7	Delaware	DE	24	Missouri	MO	40	South Dakota	SD
8	District of Columbia	DC	25	Montana	MT	41	Tennessee	TN
9	Florida	FL	26	Nebraska	NE	42	Texas	TX
10	Georgia	GA	27	Nevada	NV	43	Utah	UT
11	Idaho	ID	28	New Hampshire	NH	44	Vermont	VT
12	Illinois	IL	29	New Jersey	NJ	45	Virginia	VA
13	Indiana	IN	30	New Mexico	NM	46	Washington	WA
14	Iowa	IA	31	New York	NY	47	West Virginia	WV
15	Kansas	KS	32	North Carolina	NC	48	Wisconsin	WI
16	Kentucky	KY	33	North Dakota	ND	49	Wyoming	WY
17	Louisiana	LA						

of Section 2. For the rest of the chapter, the terms 'network' and 'graph', 'node' and 'vertex', 'link' and 'edge' are used interchangeably. They mean the same. The layout for the US States Network graph presented in Figure 1 is drawn using the Fructherman Reingold layout algorithm (Fruchterman and Reingold, 1991), available in Gephi (Cherven, 2013).

BACKGROUND

Centrality Metrics

In this section, we introduce the centrality metrics for which we will run their respective algorithms on the US States Network graph and present the results (including their distribution and ranking of the vertices). The four centrality metrics analyzed are: degree centrality, eigenvector centrality, closeness centrality and betweenness centrality. Since the US States graph is an undirected graph, the adjacency matrix of the graph is symmetric and there is only one value per vertex for each centrality metric.

Degree Centrality

The degree centrality (DegC) of a vertex is the number of edges incident on it. Table 2 presents the degree centrality of the vertices and the corresponding rank (in the decreasing order of their values) in the US States network graph; vertices with identical values for DegC have the same rank. The state of Missouri has the largest degree centrality value of 9, followed by the state of Tennessee with the second largest degree centrality value of 8. The state of Maine has the smallest degree centrality value of 1 (as New Hampshire is its only adjacent state). There are no ties among vertices for the largest and second largest values of degree centrality as well as for the vertex with

Related Content