IEEE International Conference On Recent Trends In Electronics Information Communication Technology, May 20-21, 2016, India

Degree Centrality, Eigenvector Centrality and the relation between them in Twitter

Prantik Howlader, Sudeep K.S

***Abstract*—In Social Media the directed links formed between the users, are used for the transfer of information. Based on previous research, the rate of information transfer in a social network depends on the strength of connections of the user in the network, which is measured by the centrality value. In this paper, based on data collected from Twitter, we perform an analysis of eigenvector centrality approach of finding the influential users. We investigate the variation in indegree and eigenvector centrality of users participating in a hashtag in Twitter, with respect to change in the amount of interactions. Here interactions are: tweets, mentions and replies. We also investigate the relationship between indegree and eigenvector centrality in a given hashtag. We make the following interesting observations. First, in Twitter, users with high eigenvector centrality need not be influential users. Second, in a given hashtag, there is an increase in users with both high indegree and eigenvector centrality when there are more user interactions. Here interactions are: tweets, mentions and replies, indicating both indegree and eigenvector centrality should be considered when finding influential users. Third, there is a positive correlation between indegree and eigenvector centrality.**

Keywords—Twitter, Eigenvector Centrality, Degree Centrality.

I. INTRODUCTION

In social media, a group of users transfer information among each other. This transfer of information among a group of users can be represented by a social graph, which can be used to analyse the interactions in social network. The analysis of user interactions in social network is called Social Network Analysis(SNA) [1], [2].

In a social network, the centrality value of a node represents how many connections are there from that node to other nodes [1]. There are many ways to measure centrality of a node to measure its effect on the social network: betweenness centrality, closeness centrality, degree centrality and eigenvector centrality.

Centrality metrics have been used to identify influential nodes in dynamic processes, such as opinion competition, epidemic spreading and rumor propagation [3], [4], [5], [6]. In epidemic spreading, the identified influential nodes are the source nodes from which virus spreads and nodes with high spreading capacity. Centrality has also been used to select nodes which are to be immunized when a virus is prevalent [3], [7].

Degree centrality of a node is defined as the measurement of number of links the node has in a network [8], [9], [10]. Eigenvector centrality measure of a node depends on the centrality score of its adjacent nodes. The rationale behind this is that a student's popularity increases if he is voted as popular by other popular students. Thus eigenvector centrality of a node depends on the centrality of the adjacent nodes, not on the number of adjacent nodes [11], [12]. Previous research shows high eigenvector centrality value of a user in Twitter corresponds to higher influencing capability of the node without taking into consideration the effect of interactions. Here interactions are: *follows*, *replies* and *mentions* [2].

This paper investigates the eigenvector centrality approach to find the influential users in Twitter, taking into consideration the interactions. Here interactions are: follows, replies and mentions of the users. Further we investigate the variation in indegree and eigenvector centrality of users participating in a hashtag in Twitter, with respect to change in amount of interactions. Here interactions are: tweets, mentions and replies. We also investigate the relation between indegree and eigenvector centrality by finding the Pearson Correlation Coefficient.

II. METHODOLOGY

A. Degree Centrality

Degree Centrality of a graph is defined as[11]: for a graph $G = (V, E)$, where $|V|$ is the number of vertices in the graph. Degree Centrality of node v_i is defined as:

$$C_{D(v_i)} = deg(v_i) \tag{1}$$

Where v_i in V. It is based on the idea that the number of "direct relations" of a user gives an indication of the structural importance of the user.

Finally, $\deg^{+(v_i)}$ gives the outdegree and $\deg^{-(v_i)}$ gives the indegree of the node v_i. Two users with same indegree, suggests that they have equal social status. While if one user X has higher indegree than user Y, then it means that user X is a celebrity compared to user Y [13].

B. Eigenvector Centrality

Eigenvector Centrality [12] gives a measure of the influence of the node based on the connections of the nodes to which it is connected. Same as degree centrality, eigenvector centrality favors nodes that have highest number of links. But unlike degree centrality, it also factors in the centrality of the adjacent nodes.It is defined as [2]:

For a given graph $G = (V, E)$, where $|V|$ is the number of vertices in the graph. Let $A = (a_{\{v,t\}})$ be the adjacency matrix, with $a_{\{v,t\}} = 1$ if vertex v and vertex t are adjacent in the graph otherwise $a_{\{v,t\}} = 0$. Eigenvector Centrality x_i of node i is defined as:

$$x_i = \mu \sum_{j=1}^{n} a_{ij} x_j \tag{2}$$