Abstract of the Ph.D. Work

Name of the Research Scholar	**:**	Chandni Saxena
Supervisors	**:**	Prof. M.N. Doja & Prof. Tanvir Ahmad
Department	**:**	Department of Computer Engineering, Faculty of Engineering and Technology, Jamia Millia Islamia, New Delhi

Title	**:**	A Data Mining Framework for Ranking of Nodes in Complex Networks

Complex networks from the real life complex systems have heterogeneity at various levels of their constitution. Ordinary homogeneous networks can be contrasted with heterogeneous networks in different aspects and this heterogeneity results in numerous descriptive properties for both nodes and links. A cardinal theme for research in complex networks takes an important and widely explored investigation to rank the potential of nodes which can play vital roles for various objectives. Structural configuration describes the underlying architecture of the network and functional properties explain processes on networks in the form of propagation such as information, disease, malware, opinion and viral marketing. In order to better understand the importance of nodes we have examined topologically-driven features of nodes made up of its strength in spreading and restraining contagion.

This PhD thesis confers the results of our studies associated with networks node ranking for the objectives of speeding propagation on networks for maximal

spreading and restraining the proceeding of misinformation upon immunization. The location based approaches to find dominant nodes for spreading influence in the network rely upon highly connected regions also called as the core of the graph. The generalized concept of k-core method is not optimal for collective influence due to overlap of influence spheres. With the target to provide distributed rankings based on k-core score, we propose connected coreness as centrality for node aiming information spreading. The Markov chain epidemiological aspect suggests that spreading power of a node can be rightly quantified by summarizing the affected nodes when infect transfers from the seed node in a completely susceptible network. On the same analogy we have considered the local network of a node to summarize the effect of its infect and propose an entropy-based information dissemination measure for finding influential nodes. We have further provided a framework to categorize the local information based centrality measures. We have defined two noise models to create modified networks topology keeping degree distribution unchanged and define numerical values for the stability of these measures.

Furthermore, for the target problem of node immunization we aimed to identify key nodes based on group centrality. The group centrality values the predetermined groups, although to provide individual scores to the nodes we have applied Shapley value payoff division. It utilizes positional power and functional influence of nodes to assign them rank scores according to Shapley-value based information delimiters. We have evaluated this measure on the SIR epidemic model. We have considered varying interaction strength in the network based on the degree of two end nodes connected by an edge and examine nonlinear infectivity (rumour) on the network to validate group based centrality measures for the problem of targeted immunization.