Chapter 8

Concluding Remarks and Future Directions

In this final chapter major contributions of the thesis are highlighted.

8.1 Summary of Contributions

Detection of accurate communities in the networks has remained utmost important in complex network analysis. In this thesis, the role of individual nodes in community formation has been studied extensively to enhance the accuracy of identified communities. Mostly community detection algorithms determine membership of a node from the viewpoint of other nodes. This work prioritized the role of individual nodes in determining membership instead of other nodes. Deepening the role of nodes, algorithms are developed for detecting both disjoint and overlapping communities. Multiple featured networks are also explored to identify communities. Designed metrics and evaluation methodologies to ensure accuracy of predicted communities. The community information has been used to develop an application. Main contributions of the thesis are listed below.

ENBC: Investigated community structure from the perspective of ego network. The notion of *mutual interest* in the relationship is introduced by utilizing personalized view of ego network. Defined two properties for communities: *Reachability* and *Isolability*. Reachability measures the ability of any nodes to reach out members of community, while Isolability accounts the ability of any community to isolate itself from rest of the network. Using these properties Ego Network Based Community (ENBC) detection algorithm is proposed.

FuzAg: Developed a Fuzzy Agglomerative (FuzAg) approach to identify both disjoint and overlapping communities. The notion of self-membership is introduced for the first time in fuzzy community detection to give equal opportunity to each node in growing community.

PSOCA: Introduced the notion of cognitive avoidance in PSO. Proposed an algorithm to identify communities in multiple featured networks. Community detection with PSOCA is more efficient than other PSO variants.

CLP: Developed a Community-based Link Prediction (CLP) algorithm as an application of community structure. Community information has been used in addition to connectivity pattern and node attributes in CLP. Various edge centrality measures are studied with CLP algorithm.

Metrics: Designed three quality metrics AVI, AVU and ANUI for better assurance of accuracy alternatively. Unlike most of the existing quality metrics that are developed based on dense connectivity, the proposed metrics are designed based on two properties of social community formation: *unification* and *isolation*. Linearity in the characteristics of AVI, AVU and ANUI can give the indication about accuracy. Such indication will be very helpful for determining accuracy especially for the networks where ground truth communities are unavailable.

RITA: Proposed RITA analysis methodology to evaluate accuracy of community detection algorithms by mitigating the trade-off between quality metrics and accuracy metrics. The RITA analysis gives an intuition about how likely an algorithm will produce accurate communities in the networks where ground truth is not available.

Visual Analysis: Developed a visual analysis method using regression line shifting mechanism to evaluate community detection and evolutionary algorithms by overcoming the difficultly arose with output variation. The methodology provides a platform to easily explain if an algorithm is performing better, actually, how good the solutions are compared to other algorithms.

8.2 Scope for Further Work

In this section, several new avenues of research that have been opened up by this thesis are mentioned below.

ENBC: Level 1.0 ego network has been used in ENBC algorithm. Higher levels of ego network such as $\eta^{1.5}$ and $\eta^{2.0}$ ego networks can also be explored for community detection. At present, ENBC can only identify disjoint communities, the notion of mutual interest can be further explored for identifying overlapping communities. Furthermore, the work can also be explored by incorporating soft computing techniques to define fuzziness in the relationship.

FuzAg: The notion of self-membership used in FuzAg algorithm can be extended for community detection in multiple featured networks and dynamic networks. The notion of self-membership can be used to develop new class of fuzzy logic covering new applications.

PSOCA: The cognitive avoidance mechanism used in PSOCA has already shown an application to community detection. This mechanism is applicable to those applications seeking optimization as well.

CLP: The usage of community information in CLP algorithm can be extended to various other centrality measures, particularly node centrality measures. Node-based centrality measures can be utilized easily, since CLP algorithm assigns community information for edges so pair of nodes can be considered by representing an edge.

Metrics: New community detection algorithm can be designed by simply optimizing AVI, AVU and ANUI. The metrics can be explored and tuned to increase the level of assurance of accuracy. At present, ANUI considers equal weightage to both Unifiability and Isolability. The analysis of the impact of assigning different weightage to both, and then development of suitable balancing mechanism would a valuable extension.

RITA: At present, RITA of all datasets have to be observed for giving an overall opinion. Therefore, development of single RITA score instead of examining different values for different accuracy weights would be great for analyzing RITA.

Visual Analysis: The regression line dominance mechanism used in the visual analysis methodology can be extended further by incorporating multiple functions.