# Towards Designing a Light Weight Social Network Analysis Tool

**Veerasingam Visithan** and **Thabotharan Kathiravelu**
University of Jaffna, Sri Lanka.
**visithan@gmail.com**

## Introduction and Motivation

Social network analysis has gained much popularity in recent years and the computer science research community has paid much attention in designing and developing newer algorithms, developing software tools and formulating theoretical foundations. There are many reasons for such increased interest. Monitoring, automated instrumenting, analyzing and evaluating the Internet or computer based social activities can help researchers and authorities to extract useful information.

Social networks are often large, complex and continuously evolving, and this poses a grater deal of challenges in developing tools for social network analysis. Information visualization can be an effective approach to help social science researchers both to explore relationships between actors and present their findings to others. For both of these purposes, making the visual representations of the social networks more understandable will be the most critical job.

## Background and Results

### Social Network

A social network consists of a number of actors or individuals connected by some kind of relationship. In Visualization , actors are represented by nodes and relationship are represented by links.

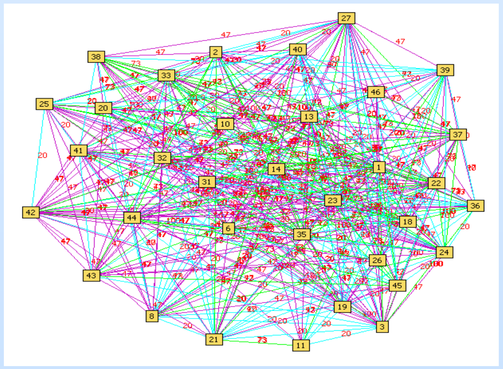

Figure 1 : Example of a Social Network

### Social Network Analysis

Social network analysis began with the detection and interpretation of patterns of social ties between actors of networks. Visualization plays an important role in Social network analysis and Social network analysis allows us to find:

- ❖ Density and degree of a network.
- ❖ Recurring patterns of connectivity.
- ❖ Clustering or community.
- ❖ Centrality of nodes.
- ❖ Connected components of a network.
- ❖ Shortest paths between nodes.

**Density and Degree**

- ❖ Density = no of edges/no of possible edges
- ❖ No of possible edges = (no of nodes-1)* no of nodes/2
- ❖ Density is useful to determine the size of a Social network.
- ❖ The Degree of a node is the number of links incident with it.
- ❖ In a social network degree indicates the popularity of a vertex.

### Visualization of a Social Network

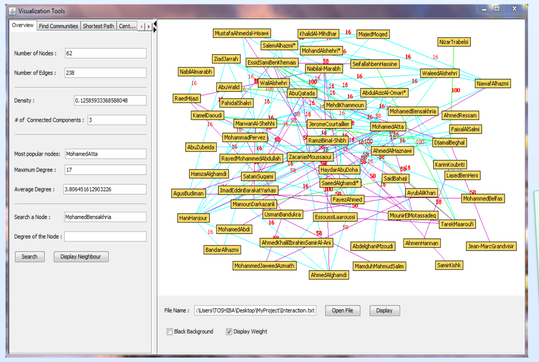

Figure 2 : Visualization of a Social Network

**Degree Centrality**

- ➢ Measures the number of direct connections that an individual node has to other nodes within a network
- ➢ Degree centralization = degree of a node/ possible maximum degree

### Network Centrality

- ❖ Identify the central/popular nodes in a network
- ❖ Approaches to measure centrality
  - Degree Centrality
  - Betweenness Centrality
  - Closeness Centrality and etc
- ❖ In Social network central node have better access to information and better opportunities to spread information

**Betweenness Centrality**

- ➢ Measures how many times a node occurs in a shortest path; measure of 'social brokerage power'.

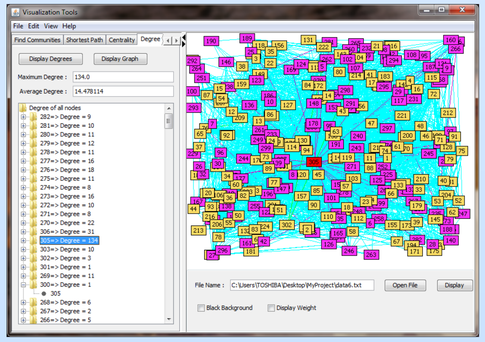

Figure 3 : Degree of a Node

### Connected Components

Identifies the distinct components of the social network. Each node belongs to exactly one components.

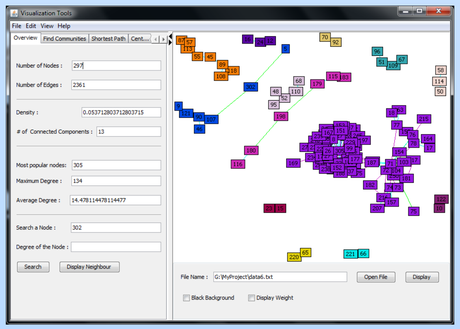

Figure 4 : Connected Components in a Social network

**Datasets**

Our dataset contains three fields. First field is the node, second is also node that is connected with the first column node by a relation and the third field is the weight of the relation.

### Shortest Paths and Path length

- ❖ A shortest path between two nodes is a minimal length path between them.
- ❖ Path length - The distances between pairs of nodes in the network.

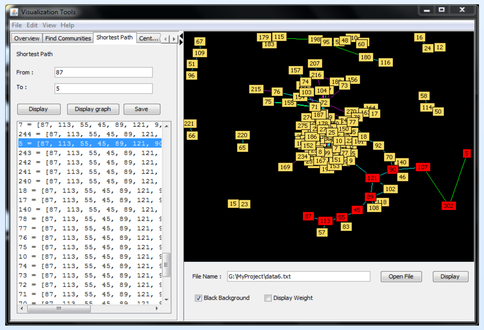

Figure 5 : Shortest path between two nodes

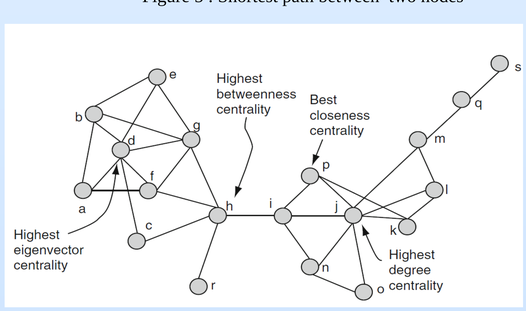


Figure 6 : Centrality in a Network

### Community or Cluster Identification

- ❖ A community within a network is a densely connected group of nodes.
- ❖ We have implemented two methods for detecting community.
  - Weak link removal method.
  - Broker nodes removal method.
- ❖ Weak link removal method removes a node that has the weight below the threshold value.
- ❖ The threshold value should be within the minimum and maximum weight of the data set.
- ❖ Broker nodes removal method removes a node that has the maximum betweenness centrality value.

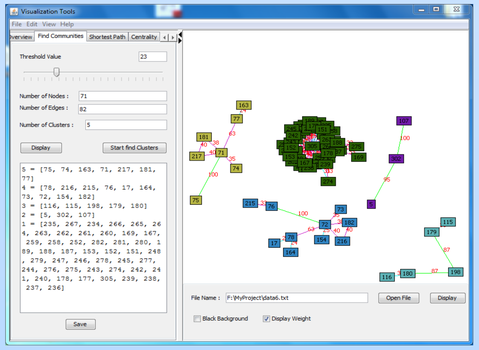

Figure 7 : Community in a Social Network

## Performance Comparison

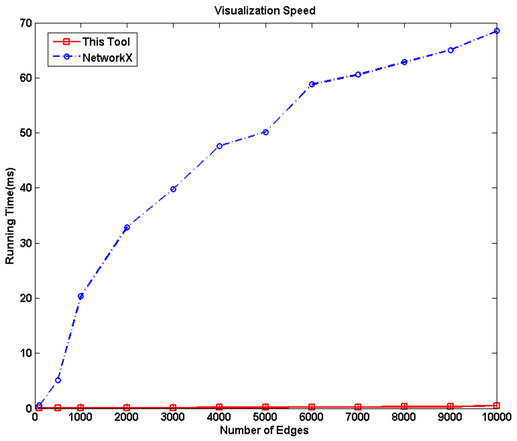


Figure 8 : Visualization Speed Comparison with NetworkX

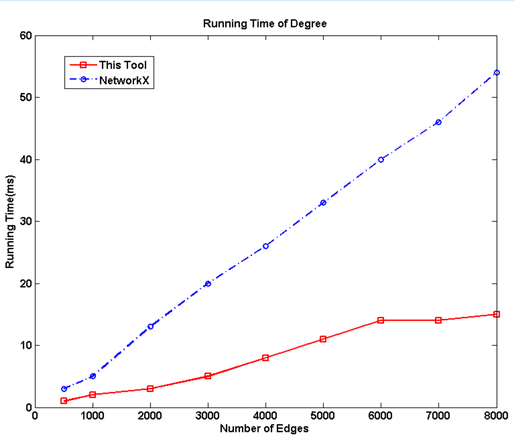


Figure 9 Degree Centrality Comparison with NetworkX

## Discussion & Conclusion

☞ In this tool, we have implemented and optimized many algorithms for social network analysis. We have also carried out a comparative study by comparing performance metrics of our tool with NetworkX.

## References

[1] A. Abraham, A. E. Hassanien, and V. Snasel. Computational Social Network Analysis. 2010.
[2] A.D. Andrea, F. Ferri, and P. Grifoni. An overview of methods for virtual social networks analysis. 2010.
[3] S. Balasubramanian and V. Mahajan. The economic leverage of the virtual community. Technical report, 2001.
[4] V. BATAGELJ. Social network analysis, large-scale. Technical report, University of Ljubljana, Ljubljana, Slovenia, 2008.
[5] U. Brandes. A faster algorithm for betweenness centrality. 2001.