Graph Algorithms for Visualizing High Dimensional Data

Abhinav Shankaranarayanan Venkataraman

Universitat Politecnica de Catalunya (UPC), Barcelona

27 June 2016

Project Research Group

The project is done under the umbrella of LARCA(Laboratory for Relational Algorithmics, Complexity and Learning) Project Directors :

- ▶ Prof. Ricard Gavalda Mestre
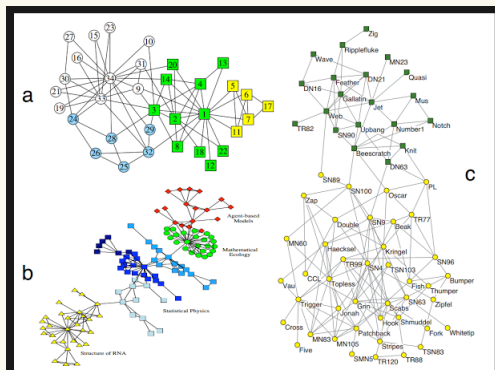- ▶ Prof. Marta Arias Vicente

What is Community?

Figure : Communities: [1]

Goal of the Project

1. To survey a few algorithms that aim in community finding keeping in mind that the input is from the medical domain.

2. To choose an algorithms that benefit the purpose of organizing graphs from medical domain and for the purpose of visualization.

3. To implement the algorithms and test the efficiency of the algorithm using variety of graphs.

4. To build a Graphic User Interface (GUI) which enables visualization of the raw input on a web browser by drawing graphs.

Planning

Planning is one of the most important part of any project. In this project we divide the project into five planning phases or stages namely,

- ▶ Required knowledge acquisition
- ▶ Paper Analysis
- ▶ Design and Implementation
- ▶ Testing I
- ▶ Testing II
- ▶ Report Writing

Economic Budget

We divide the budget into 3 major categories:

- ▶ Hardware budget
- ▶ Software Budget
- ▶ Human Resource Budget

Total budget is the sum total of the three budget.

Amortized cost : Amortized cost is that accumulated portion of the recorded cost of a fixed asset that has been charged to expense through either depreciation or amortization.

Sustainability

The project is

- ► Economically sustainable
- ► Socially sustainable
- ► Environmentally sustainable

Graph

A Graph *G* is formed by two finite sets, the set $V = \{\ v_1, v_2, \ldots, v_n\ \}$ of vertices(also called nodes) and the set $E = \{\ e_1, e_2, \ldots, e_n\ \}$ of edges where each edge is a pair of vertices from *V*, for instance,

$$e_i = (v_j, v_k)$$

is an edge from v_j to v_k represented as *G*=(*V*,*E*).

State-of-the-art in Community Detection

State-of-the-art in Visualization

Louvain Community detection Algorithm

This section has some examples of common slideshow elements.

Personal Learning

Since the project had more scope for exploration. My interest in Data Visualization has increased. My interest in graphs has increased. My python programming skill has also increased along with that I have also learned to code for web technologies on my own.

Software tools

1. git
2. github pages
3. Linux OS

List of References that were used

📄 Santo Fortunato.
Community detection in graphs.
Physics reports, 486(3):75–174, 2010.

📄 Jure Leskovec and Andrej Krevl.
SNAP Datasets: Stanford large network dataset collection.
http://snap.stanford.edu/data, June 2014.

Thank you

Thank you for all those who supported me throughout the project.
It was a Great time at Barcelona working with Prof.Ricard and
Prof.Marta.

📄 Santo Fortunato.
Community detection in graphs.
Physics reports, 486(3):75–174, 2010.

📄 Jure Leskovec and Andrej Krevl.
SNAP Datasets: Stanford large network dataset collection.
http://snap.stanford.edu/data, June 2014.