

GRAPH THEORY AND IT'S APPLICATION IN PAGE RANK

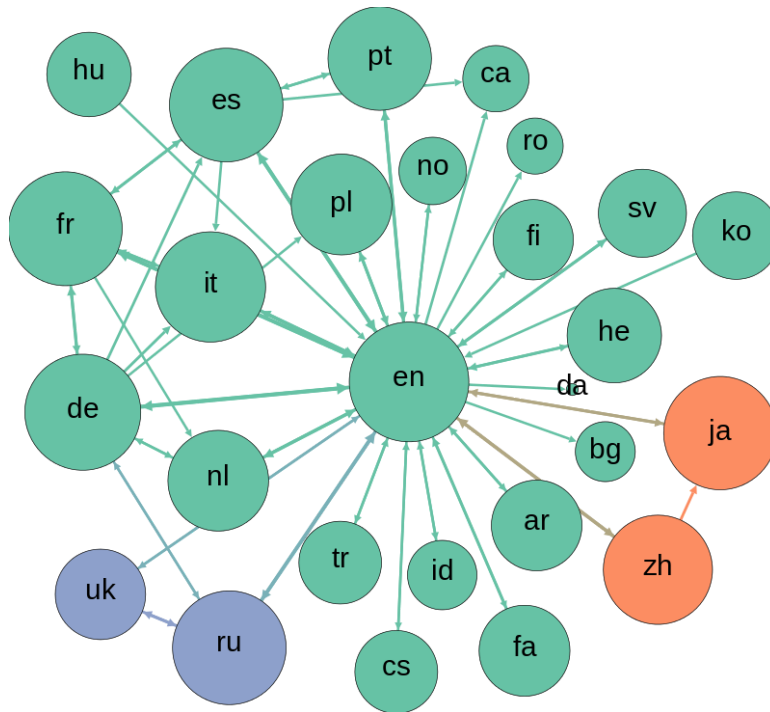
TEAM MEMBER : Pulkit Sareen ; 9929694761 ; pulkit_sareen@yahoo.co.in

INTRODUCTION

We live in a computer era. Internet is part of our everyday lives and information is only a click away. Just open your favourite search engine, like Google, AltaVista, Yahoo, type in the key words, and the search engine will display the pages relevant for your search. But how does a search engine really work? Probably using basic algebra, matrices and advanced graph theory!!

DESCRIPTION

In mathematics and computer science, graph theory is the study of graphs , which are mathematical structures used to model pairwise relations between objects.



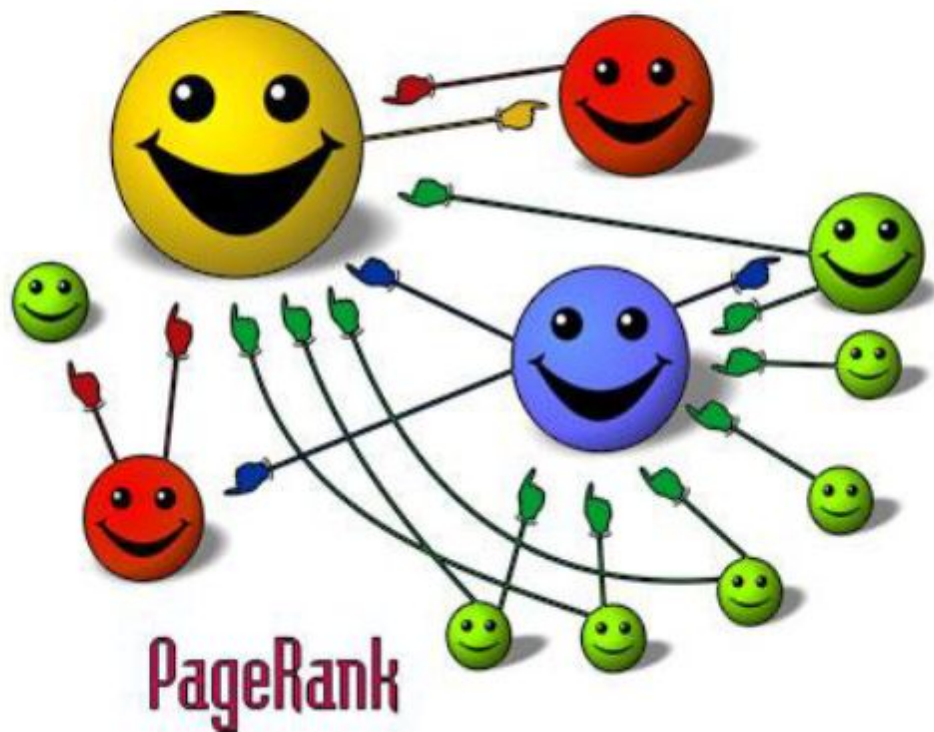
The network graph formed by Wikipedia editors (edges) contributing to different Wikipedia language versions (vertices) during one month in summer 2013.

GOOGLE

Google's PageRank algorithm, though mathematically elegant and very powerful, is very easy to understand. The first step to understanding PageRank is to view the Web as a giant graph. Each web page is a node in the graph and each hyperlink is a directed link connecting two nodes.

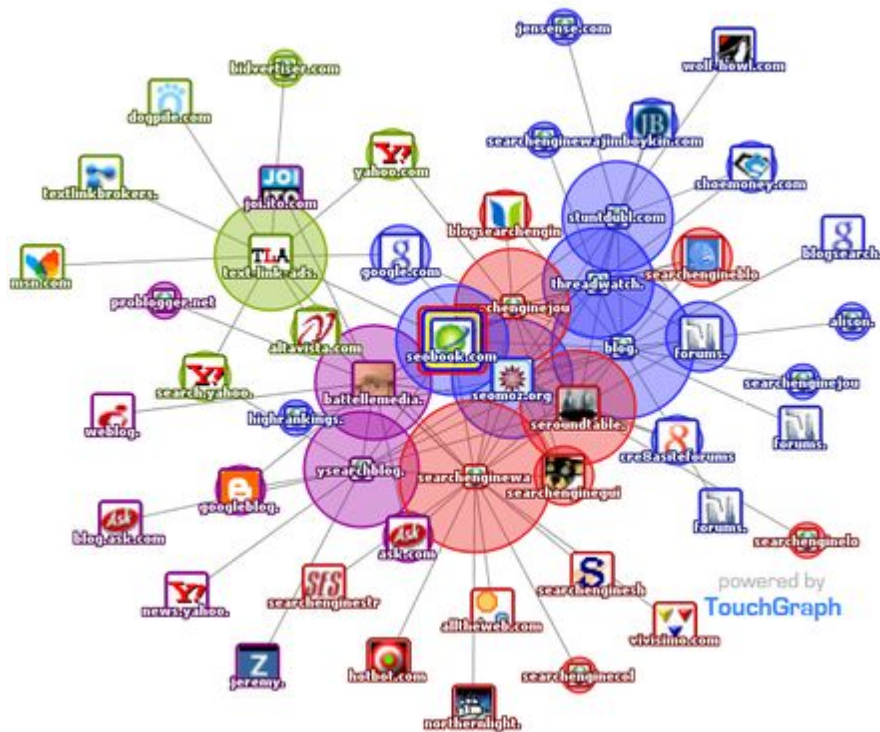
PAGERANK

Google's Page Rank



WEB

The pages and hyperlinks of the World-Wide Web may be viewed as nodes and edges in a directed graph. This graph has about a billion nodes today, several billion links, and appears to grow exponentially with time.



SKILL AND CHALLENGES

The project requires us to learn and understand the concepts of graph theory used in page ranking.

It also requires us to have basic knowledge of matrices, eigenvectors, eigen states and probability. It requires us to convert each graph in the form of a matrix.

It requires study of stochastic matrix and in-depth analysis on damping factor which is used in such matrix.

We will use this to calculate page rank using various matrix operations.

TIMELINE **

- 1) First, Second Week : Will start the study of all the theories. Mainly the graph theory. (Will require 1.5 - 2 Weeks or more.!))
- 2) Third Week : Will start studying the pagerank algorithm.
- 3) Fourth Week : Will start analysing the page rankings based on the algorithm. Analysis will be done by taking various instances with different words used (synonyms) and comparing results.
- 4) Fifth Week : Will analyse the web as a graph - assuming the graph structure in the web. As shown in the above diagram, wikipedia as a graph.

** May change as per requirement.

STUDY REFERENCES :

- 1) <http://nptel.ac.in/courses/111106050/>
- 2) Topics In Structural Graph Theory. By : Lowell W. Beineke and Robin J. Wilson
(Cambridge Publishers)
- 3) <https://en.wikipedia.org/wiki/PageRank>
- 4) <https://www.briggsby.com/methods-for-evaluating-freshness/>
- 5) https://en.wikipedia.org/wiki/Markov_chain