**Submitted by: YASH BANSAL 2014A7PS0119H, SHIVAM AGRAWAL 2014B3A70940H,**

**AMAN GUPTA 2014A7PS0201H,** **VAIBHAV AGRAWAL 2014B3A70501H**

**DESIGN DOCUMENT**

**Packages used :**

C++ - bits/stdc++.h and boost

Python – networkx, matplotlib, numpy, scipy and anaconda

We also increased the stack and heap memory limits using the pragma macros.

The formula of page rank algorithm is as follows:

**Formula :**

We implemented the code using sparse matrix representation of list of lists. This was implememnted with vector of vectors.

We also printed a graph for error at each iteration. This was Interesting as there was no monotonic nature but the errors might increase and decrease.

No. of nodes - 6301

No of Edges - 20777

**Dataset metadata is:**

# Directed graph (each unordered pair of nodes is saved once): p2p-Gnutella08.txt

# Directed Gnutella P2P network from August 8 2002

**Time Complexity:**

For calculating and printing the results in a file – 13.2 s

For visualizing – 3 s

**Pros:**

As we can see, Our pagerank algorithm works very smooth even for larger database. We have better time and space complexity.

**Cons:**

It does not use the benefit of parallel processing which is the basic fundamental for scalable solutions.