

		Finolex Academy of Management and Technology, Ratnagiri	
		Department of Information Technology	
Subject name: Big Data Lab			Subject Code: ITC801
Class	BE IT	Semester – VIII (CBSGS)	Academic year: 2019-20
Name of Student	Kazi Jawwad A Rahim		QUIZ Score :
Roll No	28	Assignment/Experiment No.	07
Title: Implementation of Page rank Algorithm			

1. Course objectives applicable: COB4. Study Page Rank in Link Analysis and concepts of Handling larger datasets
2. Course outcomes applicable: CO4- Implement use of combiners to consolidate results and ability to handle larger datasets
3. Learning Objectives: <ol style="list-style-type: none"> To understand concept of Page rank To understand teleportation factor To program Page rank computation in C/C++ To prove that teleportation helps to counteract problems caused by dead ends and spider traps
4. Practical applications of the assignment/experiment: Page rank is used by Google Search engine for indexing of webpages and giving results for search queries
5. Prerequisites: <ol style="list-style-type: none"> Understanding of Internet Technologies
6. Hardware Requirements: <ol style="list-style-type: none"> PC with 4GB RAM, 500GB HDD, 7. Software Requirements: <ol style="list-style-type: none"> Access to C/C++ compiler Internet access if online compiler is used

8. Quiz Questions (if any): (Online Exam will be taken separately batchwise, attach the certificate/ Marks obtained) <ol style="list-style-type: none"> What is a Page rank? What is Teleportation? What is a dead end? What is Random Surfer?

9. Experiment/Assignment Evaluation:			
Sr. No.	Parameters	Marks obtained	Out of
1	Technical Understanding (Assessment may be done based on Q & A <u>or</u> any other relevant method.) Teacher should mention the other method used -		6
2	Neatness/presentation		2
3	Punctuality		2
Date of performance (DOP)		Total marks obtained	10
Date of checking (DOC)		Signature of teacher	

PageRank Program with Teleportation:

Department of Information Technology

PageRank without Teleportation:

```
#include<iostream>
using namespace std;
int main ()
{
    int i, j, k,n;
    float
A[7][7]={0,0,0,0.33,0,0,0,1,0,0,0,0.5,0,0,0,1,0,0,0.5,1,0,0,0,1,0,0,0,0,0,0,0,
0,0,0.33,0,0,0,0,0,0,0.33,0,0,1};
    float B[7][1]={0.143,0.143,0.143,0.143,0.143,0.143,0.143};
    float T[7][1],C1[7][1]={0};

    for (i = 0; i < 7; i++)
    {
        for (j = 0; j < 1; j++)
        {
            C1[i][j] = 0;
            for (k = 0; k <7; k++)
            {
                C1[i][j] += A[i][k] * B[k][j];
            }
        }
    }

    cout << "\nProduct of matrices ";
    for (i = 0; i < 7; i++)
    {
        for (j = 0; j < 1; j++)
            cout << C1[i][j] << ", ";
    }

    for(n=0;n<60;n++)
    {
        for (i = 0; i < 7; i++)
        {
            for (j = 0; j < 1; j++)
            {
                T[i][j] = 0;
                for (k = 0; k <7; k++)
                {
                    T[i][j] +=(A[i][k] * C1[k][j]);

                    // T[i][j] += ((0.8*A[i][k])+0.03) * C1[k][j];
                }
            }
        }

        cout << "\nProduct of matrices ";
        for (i = 0; i < 7; i++)
        {
            for (j = 0; j < 1; j++)
            {
                cout << T[i][j] << ", ";
                C1[i][j]=T[i][j];
            }
        }
    }

    return 0;
}
```

13. Observations

1. The pagerank calculation without teleportation will cause the dead ends have maximum pagerank,
2. In Pagerank calculation with teleportation factor, the values don't converge to one rather they are reflecting actual nature of the page.
3. Total 100 Iterations were performed.

14. Results:

With Teleportation

```
Product of matrices 0.105942, 0.142553, 0.255622, 0.244656, 0.0419044, 0.105942, 0.512219,
Product of matrices 0.106854, 0.143781, 0.257823, 0.246763, 0.0422652, 0.106854, 0.51663,
Product of matrices 0.107775, 0.145019, 0.260043, 0.248888, 0.0426291, 0.107775, 0.521078,
Product of matrices 0.108703, 0.146268, 0.262283, 0.251031, 0.0429962, 0.108703, 0.525565,
Product of matrices 0.109639, 0.147527, 0.264541, 0.253192, 0.0433664, 0.109639, 0.530091,
Product of matrices 0.110583, 0.148797, 0.266819, 0.255373, 0.0437398, 0.110583, 0.534655,
Product of matrices 0.111535, 0.150079, 0.269116, 0.257572, 0.0441165, 0.111535, 0.539259,
Product of matrices 0.112495, 0.151371, 0.271434, 0.259789, 0.0444964, 0.112495, 0.543903,
Product of matrices 0.113464, 0.152674, 0.273771, 0.262026, 0.0448795, 0.113464, 0.548586,
Product of matrices 0.114441, 0.153989, 0.276128, 0.264283, 0.0452659, 0.114441, 0.55331,
Product of matrices 0.115426, 0.155315, 0.278506, 0.266558, 0.0456557, 0.115426, 0.558074,
Product of matrices 0.11642, 0.156652, 0.280904, 0.268854, 0.0460488, 0.11642, 0.56288,
Product of matrices 0.117423, 0.158001, 0.283323, 0.271169, 0.0464454, 0.117423, 0.567726,
Product of matrices 0.118434, 0.159362, 0.285762, 0.273504, 0.0468453, 0.118434, 0.572615,
Product of matrices 0.119454, 0.160734, 0.288223, 0.275859, 0.0472487, 0.119454, 0.577546,
Product of matrices 0.120482, 0.162118, 0.290705, 0.278234, 0.0476555, 0.120482, 0.582519,
Product of matrices 0.12152, 0.163514, 0.293208, 0.28063, 0.0480659, 0.12152, 0.587535,
Product of matrices 0.122566, 0.164922, 0.295733, 0.283046, 0.0484797, 0.122566, 0.592594,
Product of matrices 0.123621, 0.166342, 0.298279, 0.285483, 0.0488972, 0.123621, 0.597696,
Product of matrices 0.124686, 0.167774, 0.300848, 0.287942, 0.0493182, 0.124686, 0.602843,
Product of matrices 0.12576, 0.169219, 0.303438, 0.290421, 0.0497429, 0.12576, 0.608034,
Product of matrices 0.126842, 0.170676, 0.306051, 0.292922, 0.0501712, 0.126842, 0.61327,
```

Without Teleportation

```
Product of matrices 0.000988727, 0.00112024, 0.0023673, 0.00264496, 0, 0.000988727, 0.967893,
Product of matrices 0.000872836, 0.000988727, 0.00210897, 0.0023673, 0, 0.000872836, 0.968766,
Product of matrices 0.000781209, 0.000872836, 0.00186156, 0.00210897, 0, 0.000781209, 0.969547,
Product of matrices 0.000695959, 0.000781209, 0.00165404, 0.00186156, 0, 0.000695959, 0.970243,
Product of matrices 0.000614316, 0.000695959, 0.00147717, 0.00165404, 0, 0.000614316, 0.970858,
Product of matrices 0.000545835, 0.000614316, 0.00131027, 0.00147717, 0, 0.000545835, 0.971404,
Product of matrices 0.000487465, 0.000545835, 0.00116015, 0.00131027, 0, 0.000487465, 0.971891,
Product of matrices 0.000432391, 0.000487465, 0.0010333, 0.00116015, 0, 0.000432391, 0.972323,
Product of matrices 0.00038285, 0.000432391, 0.000919856, 0.0010333, 0, 0.00038285, 0.972706,
Product of matrices 0.000340989, 0.00038285, 0.00081524, 0.000919856, 0, 0.000340989, 0.973047,
Product of matrices 0.000303552, 0.000340989, 0.000723839, 0.00081524, 0, 0.000303552, 0.973351,
Product of matrices 0.000269029, 0.000303552, 0.000644541, 0.000723839, 0, 0.000269029, 0.97362,
Product of matrices 0.000238867, 0.000269029, 0.000572582, 0.000644541, 0, 0.000238867, 0.973859,
Product of matrices 0.000212699, 0.000238867, 0.000507896, 0.000572582, 0, 0.000212699, 0.974071,
Product of matrices 0.000188952, 0.000212699, 0.000451565, 0.000507896, 0, 0.000188952, 0.97426,
Product of matrices 0.000167606, 0.000188952, 0.000401651, 0.000451565, 0, 0.000167606, 0.974428,
Product of matrices 0.000149017, 0.000167606, 0.000356558, 0.000401651, 0, 0.000149017, 0.974577,
Product of matrices 0.000132545, 0.000149017, 0.000316622, 0.000356558, 0, 0.000132545, 0.97471,
Product of matrices 0.000117664, 0.000132545, 0.000281561, 0.000316622, 0, 0.000117664, 0.974827,
Product of matrices 0.000104485, 0.000117664, 0.000250209, 0.000281561, 0, 0.000104485, 0.974932,
Product of matrices 9.29152e-05, 0.000104485, 0.000222149, 0.000250209, 0, 9.29152e-05, 0.975025,
Product of matrices 8.25689e-05, 9.29152e-05, 0.000197401, 0.000222149, 0, 8.25689e-05, 0.975107,
```

17. References :

- [1]"Google Press Center: Fun Facts". www.google.com. Archived from the original on 2001-07-15.
- [2]"Facts about Google and Competition". Archived from the original on 4 November 2011. Retrieved 12 July 2014.
- [3] Sullivan, Danny. "What Is Google PageRank? A Guide For Searchers & Webmasters". Search Engine Land. Archived from the original on 2016-07-03.
- [4]Brin, S.; Page, L. (1998). "The anatomy of a large-scale hypertextual Web search engine" (PDF). Computer Networks and ISDN Systems. 30: 107–117. doi:10.1016/S0169-7552(98)00110-X. ISSN 0169-7552. Archived from the original on 2015-09-27.
- [5]Gyöngyi, Zoltán; Berkhin, Pavel; Garcia-Molina, Hector; Pedersen, Jan (2006), "Link spam detection based on mass estimation", Proceedings of the 32nd International Conference on Very Large Data Bases (VLDB '06, Seoul, Korea) (PDF), pp. 439–450, archived (PDF) from the original on 2014-12-03.