Name: _____ Roll #: _____ Section: _____

National University of Computer and Emerging Sciences, Lahore Campus

Course:	**Social Network Analysis**	**Course Code:**	**DS5115**
Program:	**MS(Data Science)**	**Semester:**	**Spring 2019**
Duration:	**90 Minutes**	**Total Marks:**	**50**
Paper Date:	**19-March-19**	**Weight**	**25%**
Section:	**ALL**	**Page(s):**	**5**
Exam:	**Midterm**		

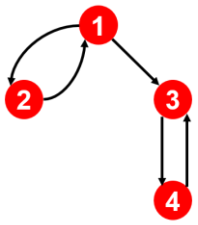
Instruction/Notes: Attempt the examination on the question paper and write concise answers. You can use extra sheet for rough work. Do not attach extra sheets used for rough with the question paper. Don't fill the table titled Questions/Marks.

Question	1	2	3	4	Total
Marks	/ 20	/ 10	/ 5	/15	/ 50

Q1) Below is equation to compute the stochastic matrix for personalized page rank where each elements in the teleport set has equal weight.

$$A_{ij} = \begin{cases} \beta\, M_{ij} + (1-\beta)/|S| & if\ i \in S \\ \beta\, M_{ij} + 0 & otherwise \end{cases}$$

Consider the following Graph



a) Compute the stochastic matrix if S={1, 2} **[5 Marks]**

b) Suppose we want to assign different weights to elements of the teleport set S = {1, 2} i.e. 20% weight to node 1 and 80 % weight to node 2. Compute the stochastic matrix. **[5 Marks]**

c) Give the rank vector for the first two iterations of page rank algorithm using the stochastic matrix computed in part (b). Show complete working [**10 Marks**]

Name: _____ Roll #: _____ Section: _____

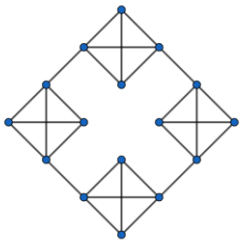
Q2) Consider a network with *N* nodes placed on a circle, so that each node connects to *m* neighbors on either side (consequently each node has degree 2m). For simplicity assume that *N* and *m* are chosen such that (*N*-1)/2*m* is an integer.

a) Calculate the average clustering coefficient $\langle C \rangle$ of this network. The clustering coefficeint of node i is $C_i = \frac{2e_i}{k_i(k_i-1)}$. Average clustering coefficient of the network is $C = \frac{1}{N}\sum_{i=1}^{N} C_i$ [**5 Marks**]

b) Calculate the average shortest path $\langle d \rangle$ of this network. $\bar{h} = \frac{1}{N(N-1)}\sum_{i,j} h_{i,j}$, where $h_{i,j}$ = is the distance from node i to node j and N(N-1) is all possible vertex pairs. [**5 Marks**]

Q3) One can think of DeepWalk as a specific case of node2vec. Explain briefly but specifically (1-2 sentences) why this is true, and how one can use node2vec to simulate DeepWalk [**5 Marks**]

Q4) Consider the graph G with 16 nodes, having 4 cliques of 4 nodes each arranged in a ring. There exists exactly one edge between any two cliques. Suppose the first phase of Louvain modularity optimization algorithm detects each clique as a single community (giving 4 communities in all). After the community aggregation phase, the new network H will have 4 nodes.



a. What is the weight of any edge between two distinct nodes in H? [**5 Marks**]

b. What is the weight of any self-edge in H? Explian your answer. [**5 Marks**]

c. What is the modularity of H (with each node in its own community)? Show complete working. [**5 Marks**]