



VII SEMESTER B.TECH. (COMPUTER SCIENCE AND ENGINEERING)

END SEMESTER EXAMINATIONS, NOV/DEC 2019

SUBJECT: GRAPH ANALYTICS FOR BIG DATA [CRA 4008]

REVISED CREDIT SYSTEM

(28/11/2019)

Time: 3 Hours

MAX. MARKS: 50

Instructions to Candidates:

- ❖ Answer **ALL** the questions.
- ❖ Missing data may be suitable assumed.

- 1A.** Write the Adjacency matrix for the following graph in figure 1A. With any one operation explain how this matrix representation is useful? What is the other way of representing this graph? Represent it. **4M**

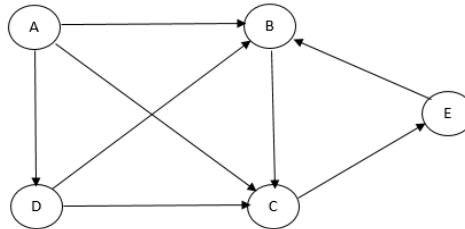


Figure 1A Directed Graph

- 1B.** How can you represent a Facebook account using a graph? Explain the concept of valence with respect to that graph. **3M**
- 1C.** What is analytics? What is the purpose of it? **3M**
- 2A.** Find out the diameter of the following graph in figure 2A by clearly showing all the steps. **2M**
- 2B.** What are two types of measures used for finding cohesive groups of communities? With the help of a graph explain each of them. **4M**
- 2C.** What is closeness centrality and betweenness centrality? Find out the betweenness-centrality of all the nodes in the graph given in Figure 2C. When betweenness-centrality is not applicable? **4M**

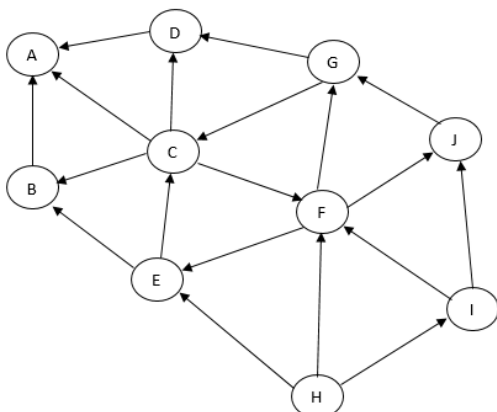


Figure 2A Directed Graph

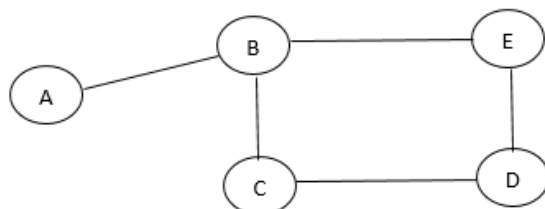


Figure 2C Undirected Graph

- 3A.** What do you mean by strongly connected and weakly connected graph? Identify whether the graph shown in question 2A is strongly or weakly connected? Justify your answer. **2M**
- 3B.** List with example the six large categories that can happen within a community? **3M**
- 3C.** Explain all the steps in detail about how Bulk Synchronous Parallel model is applied on Dijkstra's SSSP algorithm for the graph given in the figure. 3C. **5M**

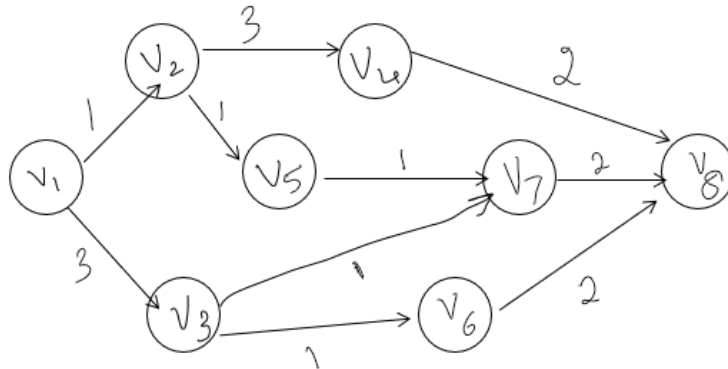


Figure. 3C Input graph for Dijkstra's algorithm

- 4A.** Write a Cypher query to find the indegree of all nodes and to find the shortest path from Danapur to New Delhi (Table 5B(b)). **4M**
- 4B.** Write a Cypher query to construct an adjacency matrix that contains 999 if there is no edge between two different nodes otherwise the weight of the edge between the nodes. Display the table under the label source, destination and weight for the graph in figure 3C. **4M**
- 4C.** Explain the different wrappers used by Giraph to interoperate with Hadoop and Relational data sources. **2M**
- 5A.** Write a Cypher query to return a graph that does not contain the immediate neighborhood of V3 in figure 3C. **4M**
- 5B.** Write a GraphX query to create a graph trainGraph, from a list of vertices i.e., the railway station names given in table 5B(a) & a list of edges that will make up the train service from table 5B(b). **3M**

Table 5B(a)

Station ID	Station Name
1	PUNE JN
2	NEW DELHI
3	PURI
4	BANGALORE CITY JN
5	DANAPUR
6	HOWRAH JN
7	JAMMU TAWI
8	DARBHANGA JN

Table 5B(b)

Train Number	Source station	Destination station
11177	PUNE JN	JAMMU TAWI
12888	NEW DELHI	PURI
12800	PURI	NEW DELHI
12200	BANGALORE CITY JN	DANAPUR
12134	PUNE JN	HOWRAH JN
12211	DANAPUR	BANGALORE CITY JN
12136	HOWRAH JN	PUNE JN
12347	BANGALORE CITY JN	NEW DELHI
12576	NEW DELHI	DARBHANGA JN

- 5C.** Write all the necessary GraphX queries to create and join the station information given in table 5C to trainGraph graph created in question 5B. **3M**

Table 5C

Station ID	City	Code	Type
1	Pune	PUNE	Cosmo
2	New Delhi	NDLS	Cosmo
3	Puri	PURI	NA
4	Bangalore	SBC	Cosmo
5	Danapur	DNR	NA
6	Howrah	HWH	Metro
9	Ahmedabad	ADI	Cosmo