

# Question Paper



## MANIPAL UNIVERSITY

### SCHOOL OF INFORMATION SCIENCES SECOND SEMESTER MASTER OF ENGINEERING - ME (Big Data and Data Analytics)

DEGREE EXAMINATION - NOVEMBER 2017

DATE : Monday, November 20, 2017

TIME : 10:00AM - 1:00PM

Text Retrieval and Search Engines [BDA 616.1]

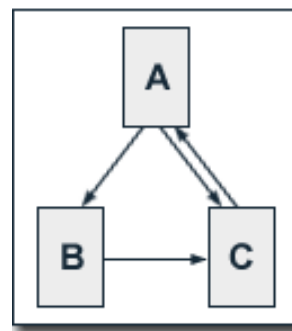
Marks: 100

Duration: 180 mins.

### ELECTIVE 2

**Answer all the questions.**

- 1) Describe the characteristics of the Web. (10)
- 2) How do you estimate the size of the index of a search engine? (10)
- 3) Explain Single-pass in-memory Indexing. (10)
- 4) What is Distributed indexing? Why is such an indexing technique required? (10)
- 5) What are the difference document parsing techniques used before creating an inverted index? (10)
- 6) Briefly explain the Porter Stemmer. (10)
- 7) Calculate the Page Rank for given graph with the teleportation factor  $\alpha=0.5$  (10)



- 8) Briefly explain the steps involved in creating an inverted index with an example. (10)
- 9) Consider these documents: (10)
  - Doc 1** breakthrough drug for schizophrenia
  - Doc 2** new schizophrenia drug
  - Doc 3** new approach for treatment of schizophrenia
  - Doc 4** new hopes for schizophrenia patientsDraw the term-document incidence matrix for this

document collection.

10) What is the Vector Space Model? (10)

-----End-----