JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- APRIL 2018

M.Tech IV Semester

COURSE CODE: 11M1WCI431   MAX. MARKS: 25

COURSE NAME: Advanced Web Mining

COURSE CREDITS: 03   MAX. TIME: 1Hr 30 Min

Note: All questions are compulsory. Carrying of mobile phone during examinations will be
treated as case of unfair means.

1.   a) What is boolean indexing in inverted indexing?
      b) How skip list is used to search occurrence faster?
      c) What is stop-press indexing? Explain the key word search under stop-press indexing.
      d) Describe map-reduce algorithm for constructing large scale inverted index

                                                                    [1+2+3+4]

2.   a) Write the Page-rank algorithm and derive its matrix equation.
      b) Compute the rank of hypertext documents in a web graph with at least 5 vertices and
      12 edges (create any) using HITS algorithm.   [4+7]

3. Write short notes on the following: [2+2]
      a) TFIDF
      b) Cosine similarity