

# CSCI5070 Advanced Topics in Social Computing

## Assignment 2

Deadline: 15:59:59, November 6 (Tuesday), 2012

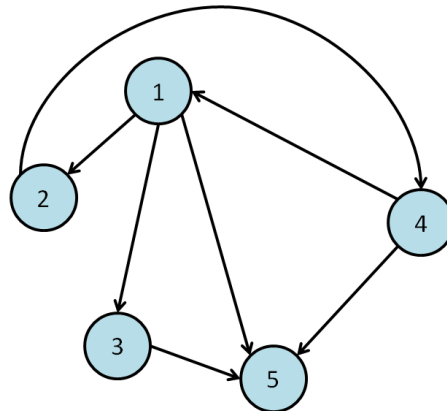
### 1. Ford-Fulkerson Algorithm.

Consider the network flow problem with the following edge capacities,  $c(u, v)$  for edge  $(u, v)$ :  $c(s, 2) = 2$ ,  $c(s, 3) = 13$ ,  $c(2, 5) = 12$ ,  $c(2, 4) = 10$ ,  $c(3, 4) = 5$ ,  $c(3, 7) = 6$ ,  $c(4, 5) = 1$ ,  $c(4, 6) = 1$ ,  $c(6, 5) = 2$ ,  $c(6, 7) = 3$ ,  $c(5, t) = 6$ ,  $c(7, t) = 2$ .

- (1) Draw the network.
- (2) Run the Ford-Fulkerson algorithm to find the maximum flow. Show each residual graph.
- (3) Show the minimum cut.

### 2. PageRank and HITS.

The the link structure of five web pages is shown in the following figure.



- (1) Suppose  $d = 0.7$ , please calculate PageRank score of each state in the first and second iterations. The initiate score of each state is 0.2.
- (2) The initialization of hub score and authority score for each node are both 0.2. Please calculate the hub and authority scores of each state in the first and second iterations.

- (3) Please try to find relevant materials to prove the convergence of PageRank algorithm (Hints: PageRank is a special case of Markov Process, where relevant theorems can be found.).

3. Memory-based Collaborative Filtering.

	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$
$u_1$	0	2	5	3	1	0
$u_2$	3	3	4	3	0	2
$u_3$	3	0	0	5	2	2
$u_4$	5	0	4	4	5	3
$u_5$	2	3	3	0	2	2

The above table shows the ratings of 5 users on 6 items (The value 0 means the user has not rated the item). Please utilize Pearson Correlation Coefficient (PCC) similarity and Memory-based CF algorithms introduced in the lecture notes to

- (1) find top 2 most similar users of  $u_3$  and estimate  $u_3$ 's rating on  $i_2$  using user-based CF.
- (2) find top 2 most similar items of  $i_5$  and estimate  $u_2$ 's rating on  $i_5$  using item-based CF.