SOCIAL NETWORKS - NPTEL July 2024

Week 6

1. In the context of a web graph, which of the following statements best describes the significance of directed links between web pages?

   (a) Directed links indicate the physical location of web pages on a server.

   (b) Directed links represent the number of visits each web page receives.

   (c) Directed links show the navigation paths from one web page to another.

   (d) Directed links determine the color and layout of web pages.

   Answer: (c)

2. Consider the implementation of Page Rank using Random Walk. Which of the following is TRUE?

   Statement I - Random walk points is higher for the nodes with higher number of inlinks
   Statement II - Random walk point is decreased by one when we reach a node

   (a) I only

   (b) II only

   (c) Both I and II

   (d) Neither I nor II

   Answer: (a)

   Explanation:

   Statement I: Random walk points are higher for the nodes with higher number of inlinks. This is true. In the PageRank algorithm, a node's rank is influenced by the number of inlinks (incoming links) it has. Nodes with more inlinks are more likely to be visited during a random walk, thereby accumulating more "points" or a higher rank.

   Statement II: Random walk point is decreased by one when we reach a node.

   This is not true. When a random walk reaches a node, the point for that node is not decreased by one. Instead, when a random walk reaches a node, the random walk point for that node is increased because it indicates another visit to that node, contributing to its overall rank.

3. In a game where participants Alex, Sam, and Jamie distribute gold coins equally, it is observed that the game converges. At the convergence state, which of the following statements is most accurate?

   (a) Alex, Sam, and Jamie always end up with an equal number of gold coins.

   (b) Alex, Sam, and Jamie may or may not have an equal number of gold coins.

   (c) The total number of gold coins with Alex, Sam, and Jamie increases as the game converges.

   (d) The game never converges if Alex, Sam, and Jamie distribute the gold coins equally.

   Answer: (b)

   Explanation: When people distribute gold coins equally, the game converges. At the convergence state, everybody might/might not have the same number of gold coins.

4. Which type of node could potentially cause an issue in the random walk (drop) gold coins distribution game?

   (a) node with highest indegree

   (b) node with least indegree

   (c) node with highest outdegree

(d) node with zero outdegree

Answer: (d)

Explanation: A node with zero outdegree has no outgoing edges, meaning it cannot transfer gold coins to any other node. This creates a significant problem in the random walk (drop) gold coins distribution game. When gold coins reach a node with zero outdegree, they become trapped there since the node has no way to pass them on to other nodes. This accumulation of coins at a single node disrupts the intended continuous flow and redistribution of coins among all nodes. As a result, the game may fail to achieve a balanced or stable distribution of coins, preventing proper convergence.

5. Examine the graph depicted in the following figure. Each circle contains a number representing the amount of gold coins the corresponding node holds. Determine the number of gold coins each node will have after one iteration of the equal sharing gold coins game.

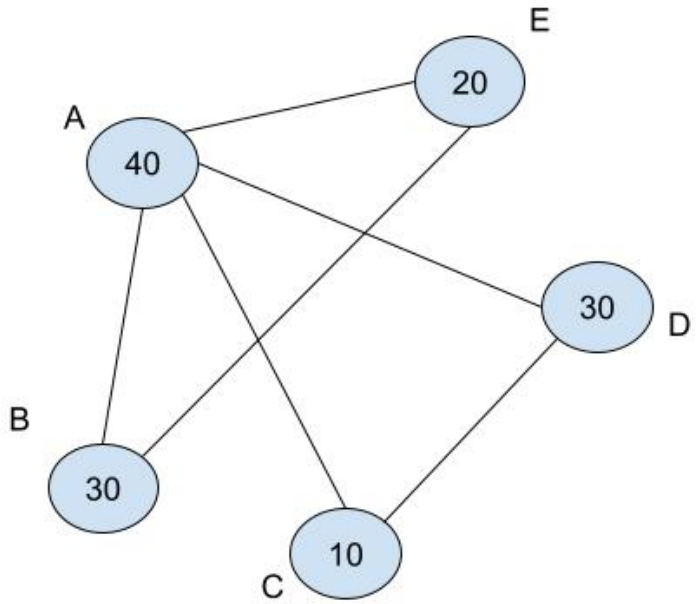

Figure 1: Network H

(a) A: 45, B: 20, C: 25, D: 15, E: 25

(b) A: 45, B: 20, C: 25, D: 15, E: 5

(c) A: 45, B: 25, C: 25, D: 15, E: 25

(d) A: 45, B: 20, C: 25, D: 25, E: 25

(e) A: 45, B: 20, C: 25, D: 15, E: 20

Answer: (a)

Explanation:

P'(A) = P(B)/2 + P(C)/2 + P(D)/2 + P(E)/2 = 15 + 5 + 15 + 10 = 45

P'(B) = P(A)/4 + P(E)/2 = 10 + 10 = 20

P'(c) = P(A)/4 + P(D)/2 = 10 + 15 = 25

P'(D) = P(C)/2 + P(A)/4 = 5 + 10 = 15

P'(E) = P(A)/4 + P(B)/2 = 10 + 15 = 25

6. Which of the following best describes the primary concept behind the Google PageRank algorithm?

   (a) The algorithm ranks pages based solely on the relevance of their content to the search query, using keywords and semantic analysis

   (b) The algorithm ranks pages by evaluating the quantity and quality of backlinks, distributing rank scores iteratively based on link structure and transition probabilities

   (c) The algorithm ranks pages by analyzing user engagement metrics such as click-through rates, time spent on page, and bounce rates

   (d) The algorithm ranks pages based on their presence and influence on social media platforms, considering factors like likes, shares, and comments

   Answer: (b)

   Explanation:

   Option (a) is incorrect because PageRank does not rank pages solely based on content relevance.

   Option (b) is correct as it accurately describes the primary concept behind PageRank, which involves analyzing backlinks and iteratively distributing rank scores based on link structure and transition probabilities.

   Option (c) is incorrect because PageRank does not incorporate user engagement metrics as its primary ranking factor.

   Option (d) is incorrect because PageRank does not rank pages based on social media influence

7. What is the relationship between indegree and page rank values for the mentioned citation network?

   (a) Indegree is not correlated with page rank values

   (b) indegree is directly proportional to page rank

   (c) indegree is inversely proportional to page rank values

   (d) None of the above

   Answer: (a)

   Explanation: While indegree (the number of incoming links to a node) can influence the PageRank of a node, it is not the sole determinant of PageRank. PageRank values are calculated based on the probability distribution of a random walker visiting nodes, which depends not only on the number of incoming links but also on the quality and rank of the linking nodes. Additionally, PageRank considers the distribution of rank scores through outgoing links.

   Thus, even though there can be a correlation between indegree and PageRank (nodes with higher indegree often tend to have higher PageRank), this correlation is not strict. High indegree does not guarantee high PageRank if the incoming links are from low-ranked nodes, and low indegree does not necessarily mean low PageRank if the incoming links are from highly ranked nodes. Therefore, indegree and PageRank values are not perfectly correlated.

8. Which of the following correctly depicts teleportation?

   (a) Jumping from the current node to its neighbor's neighbor.

   (b) Going back to the previous node which was explored.

   (c) Jumping to any random node in the network

   (d) Jumping to the node in the network which has maximum outdegree.

   Answer: (c)

   Explanation: It has been shown in the lecture videos that once the random walk algorithm gets stuck/trapped at a node, it jumps to a randomly chosen node in the network. This concept is known as Teleportation.

9. What happens if one or more nodes have no outlink while implementing the page rank algorithm?

   (a) Points in every node in the network keeps increasing without bounds

   (b) Points get accumulated in such nodes with no outlinks

(c) Points get to reach zero in the nodes with zero outlinks

(d) Points in every node in the network converges in exactly one iteration

Answer: (b) Explanation: In the PageRank algorithm, nodes with no outlinks are referred to as "dangling nodes." When a random walker reaches a dangling node, it has nowhere to go.This causes accumulation of points in nodes with no outlinks.

10. Identify the correct sequence of implementing page rank using points distribution method repeating this process until convergence.

I. Assign 100 points to each of the n nodes

II. Get ranking based on the points accumulated

III. Create a directed graph with n nodes

IV. Distribute points

(a) I, II, IV, III

(b) II, I, IV, III

(c) III, I, IV, II

(d) III, II, I, IV

Answer: (c)

Explanation: The correct sequence for implementing the PageRank algorithm using the points distribution method involves:

Creating the directed graph (Step III): This is the initial step where the structure of the network is defined, including the nodes and the directed edges between them.

Assigning initial points (Step I): Each node is given an initial value of 100 points (or any arbitrary value) to start the process.

Distributing points (Step IV): Points are distributed from each node to its outlinks according to the algorithm, simulating the random walk process. This step is repeated iteratively.

Getting the ranking (Step II): After points are distributed and the system reaches convergence (where the points distribution stabilizes), the final ranking of the nodes is determined based on the accumulated points.