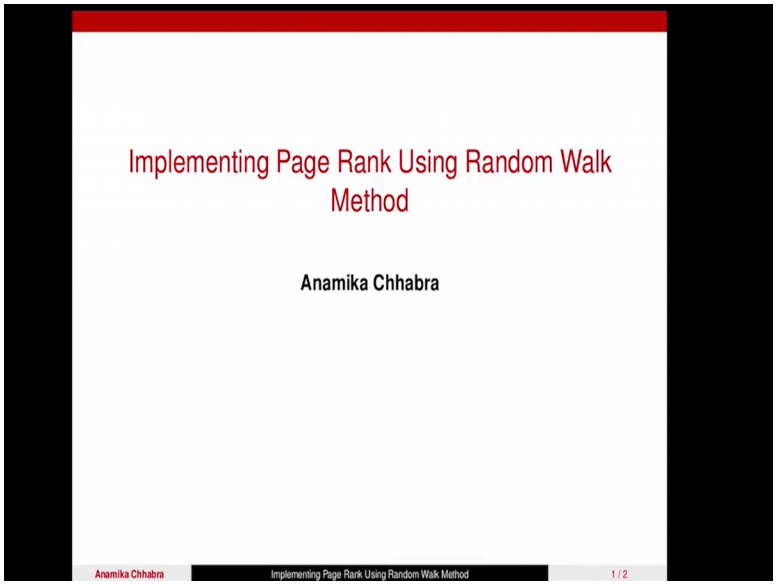
Social Networks
Prof. S.R.S. Iyengar
Prof. Anamika Chhabra
Department of Computer Science
Indian Institute of Technology, Ropar

Link Analysis
Lecture – 84
Implementing PageRank Using Random Walk Method – 1

(Refer Slide Time: 00:05)



Hey everyone, in the previous video we implemented page rank using points distribution method. In this video we are going to Implement Page Rank Using Random Walk Method. Now what exactly is random walk method? So, basically we are going to choose 1 node from the network uniformly at random. And, we will look at its neighbors; we will choose one neighbor out of them uniformly at random. And then we will look at the neighbors of that node, and we will keep doing this. And that is how we will randomly walk on the network that precisely what random walk is.

Now, there are two things that we are going to do along with this walk. And that is first thing that we will do is as and when we choose a node we will increment its points by 1. Since we used points in the previous method as well which are different sort of points, let us call these points as random walk points just to differentiate things.

So, as I was saying as and when we reach a node we increment its random walk points by 1 ok. That is one thing that we do, second thing is while we are randomly walking on the network it might so happen that we reach a node which has no out links; which has no neighbors where we can go right. So, in that case what we do is we out of all the nodes in the network we choose one uniformly at random and we resume ours our random walk from that node.

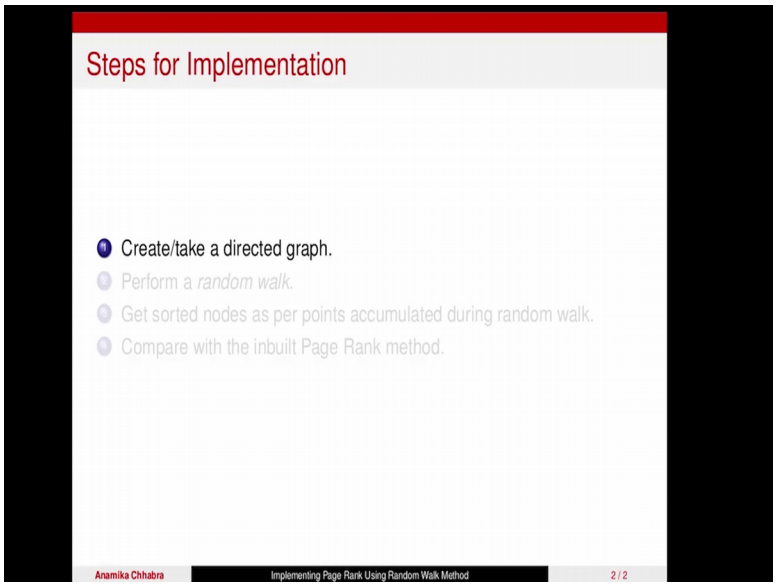
Now this is precisely called teleportation, which is a frequently used term when you talk about page rank and random walks. So, teleportation is basically choosing a node uniformly at random from the network and starting our basically resuming our random walk from that point.

Now, what is the basic idea behind this concept? How is it I mean it is, it looks counterintuitive that we are trying to implement page rank using random walk; it sort of does not make sense intuitively. So, let me tell you the basic idea that is used here. As we randomly walk on the network, a node which has a number of huge number of in links. There is a high probability that we will reach at this node more often right. And anytime we reach a node we are incrementing its random walk points.

So, the number of times we reach a node we are keeping a track of that. And a node which has high number of in links we will be reaching that node quite often. On contrary if we have a few nodes which have very less number of in links, we will be reaching them very less often and that is what is recorded by the number of random walk points. So, that is a basic idea because in page rank also the idea is same.

If lot of nodes point to you then you are in important ok. And another thing that it captures is if a lot of important nodes points to you then you are important. So, in the random walk also that thing will be captured that if a lot of nodes, which are reached quiet often and they are pointing to you, you will be reached quiet often. And, the same thing is captured by the count that we are keeping random walk points that we are keeping. So and based on those points we are going to rank the nodes; which will be precisely same as a ranking that we get from the page rank technique. So, let us see the steps that we are going to implement ok.

(Refer Slide Time: 03:54)



As in the previous case we are going to take a directed graph. So, we can either take it using some generator from network x function or we can create it ourselves. So, we created in the previous video we are just going to make use of that code again.

Second step that has to be implemented is the main step which is performing a random walk. As I just explained to you, we are going to implement those steps. After performing a random walk we are going to have count of the number of times we reached a node that is the random walk points. So, we are going to rank the node based on the random walk points that we obtained in the previous step.

In the last step as we did in the previous video as well we are going to compare a results, with the results that we obtained from the inbuilt page rank method from network x. So, these are the sequence of steps that we are going to follow; first, third and fourth steps are the same more or less same as in the previous video. So, the main thing that has to be implemented is the second one that is the random walk which we are going to in the next video.