

PAGE RANK: HOW DOES GOOGLE WORK? 03

You know what I am going to do with this data. I now have a what is called a network and my network look like this do you see the nodes here and the edges, look at this arrow from this node to this node it just represent that this person found this person impressive. If A finds B impressive B may not find A impressive correct? Now my question is how so you find the most impressive person here in this network and obvious answer that we will have is ok at that person who has a lot of arrows coming in by that we mean many people found him impressive a number of people finding someone impressive should not be an indication of how impressive a person is. For example assume hundred friends of mine says I am a nice person that is different from the prime minister of India saying "I am a nice person" in one of his speeches. Right there is a huge difference so what do you infer? It is not how many people who elect who vote for you it's about who are those people who vote for you, maybe we should consider that right? So now on this network I would like to ask this question who is the most impressive person? And I am going to give you a technique to find it. You may not even see why exactly we are using this technique to find the most impressive person but with time it will be very clear. Here goes my algorithm, here goes my technique what I will do is on this graph that you are seeing the network I will start from some node and take what is called a drunkards' walk also called the random walk the point is you start from the node let's say this node you have so many lines going from this node forget about what this network represents I am going to play around this network. I am going to jump to one of the lines and land at the other end of this line the point whatever you are seeing I am going to land there but I have so many choices to take I will pick one of them randomly then I jump there once I jump there I mark one here saying that I have seen it already and from here I will jump to another node again from here there are many lines going out I will choice one of them randomly and jump there and call it I have visited it once I keep doing this I keep doing this if I visited node more than once I will make a note of it I will call it I visited this node twice and so on and finally I will do this lets say a million times may if I am not happy I will do it a couple of million times or even more than that and after doing this a million times I will see what is the points collected by individual nodes. And I will look at that node which has collected the maximum number of points and I will declare that node as the winner. Why do I do this? What does this signify? You will get to know very soon.