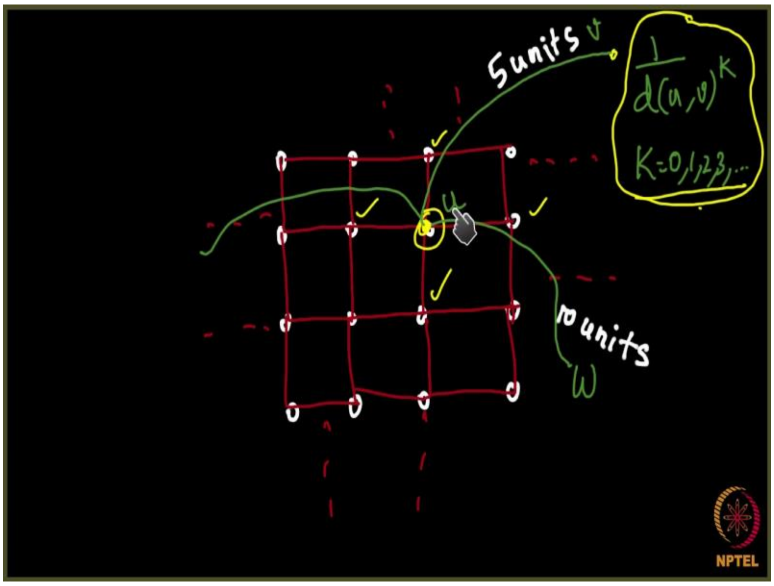
Social Networks
Prof. S. R. S. Iyengar
Department of Computer Science
Indian Institute of Technology, Ropar

The Small World Effect
Lecture - 149
Decentralized Search – III

(Refer Slide Time: 00:05)



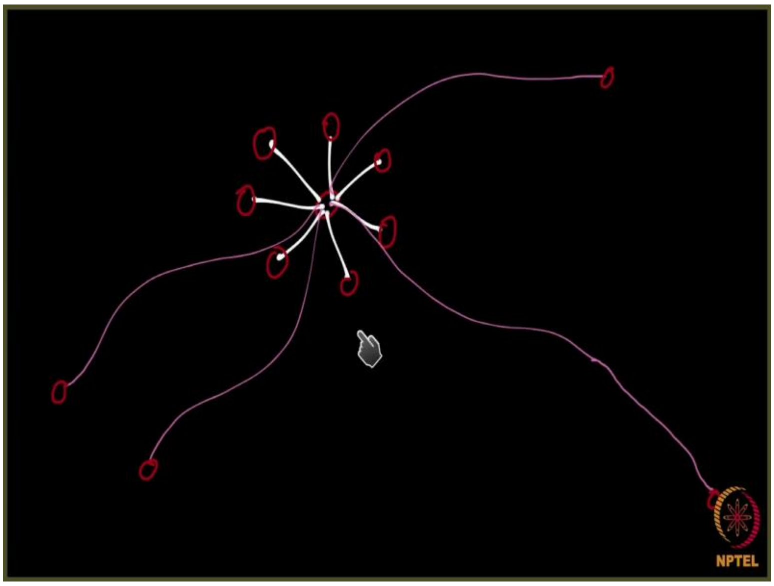
So, K here equals 1 2 0 1 2 3 here denotes that you are rewiring, basically your friendship that u maintains with people who are not in his neighbourhood. The friendship that u maintains is sort of it is based on the distance of that node from u; further away the node less probable is a use u being friends with them. For example, let us say the this is a node which is some 10 units 10 distance away from u and this one is some 20 distance away from u.

So, this is less probable, and this is more probable ok, that is what this distance function captures and for a fixed K. So, what do you mean by this? If the distance is let us say 5 here let me let me write that down. So, if this distance is let us say 5 units and this one is 10 units, the probability of u being friends with v is more than the probability of u being friends with this person ok. Let me give him some name, let me call him w this person let its call let us say this node is called w ok.

The probability of u being friends with v is a lot more because, it is closer just 5 unit away; 5 units I mean 5 edges away; it is not just 1 edge ok. In case u and v are 5 edges away then you put an edge between u and v based on a probability function ok. What is the probability function here? Nothing it is very simple you judge whether you want to put an edge or not by looking at the distance from u. This is far away from u this is 10 units far away from u. This is 5 units closer to u, this is more probable to become a neighbour of u, v is more probable to become neighbour of u than w.

What decides this? It is this equation that decides this. So, 1 over if you let us say assume I K equals 2 then $1/5^2$ is the probability of u being adjacent to v. And, $1/10^2$ is the probability of u being adjacent to w that is if K equals 2. If K equals 1 then 1/5 is a probability of v being adjacent to u and 1 over 10 is a probability of w being adjacent to u ok. Let us not worry much about this fact, let us only observe the following all I am saying here is there is a node here. And he is of course, friends with his immediate neighbours, a few of his immediate neighbours alright; very clear till here.

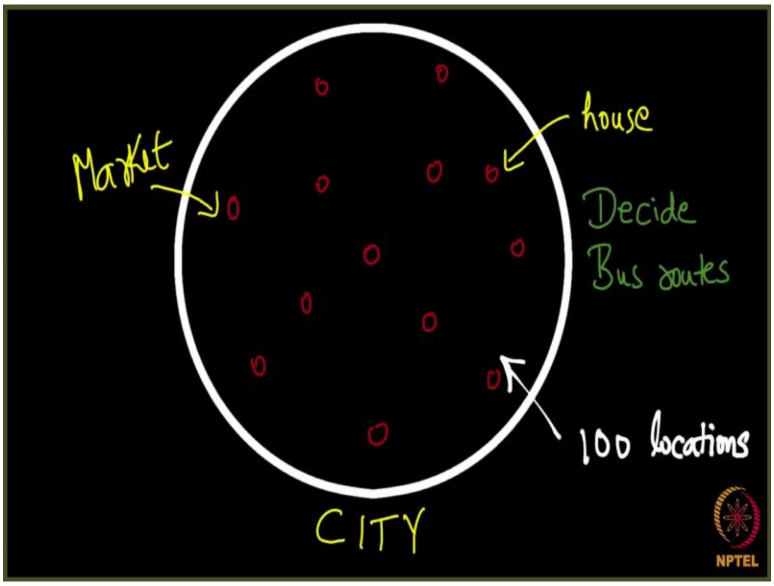
(Refer Slide Time: 03:09)



And all that I was saying before is that he also takes part in being friends with far away people, far away people let us say a few of them. Few friends are far away people for him, but then that friendship simply depends on the distance ok. By that I mean further away they are less probable that this person is friends with him. Who decides it? I am going to decide it based on the inverse of the distance. So, if let us say in plain English language if this is some 5 kilometres away.

This is 20 kilometres away, the probability of friendship here is rare probability of friendship here is a more probable ok. So, I once I give you a nice example you will realise what is the happening here, what is the motivation for this distance function alright. So, here goes the right motivation; assume there is a city you are a city planner.

(Refer Slide Time: 04:27)

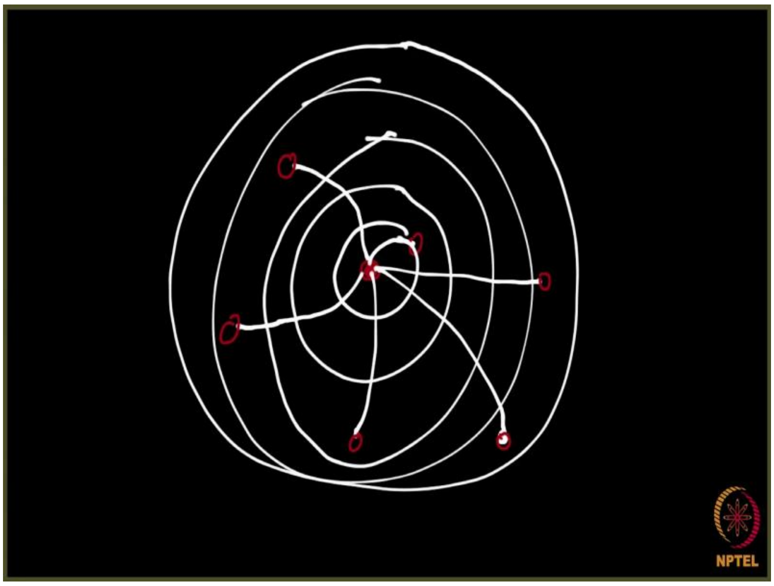


Let us assume you are a city plan, and this is the city and here is the centre of the city alright and you may want to go to different places here. These are all different locations, for the time being just forget the graph here. We are not talking about a graph; we are just talking about let us say we are planning a newly built city. The city is newly built, I am planning some road networks here, public transport here to be precise; planning let us plan some public transport here. So, now assume this is your house and this is some market a supermarket or a shopping complex. So, whatever you want, now you want to go from here from your house, you would like to go to market.

Of course, you can take your car and then go, but the point is you would like to use public transport. There are so many locations here, as you can see there are let us say 100's of locations here. You cannot basically put a bus a public transport bus from any node to any other node here; from your home to market you may not have a bus. So, let us say you may want to go from your home to let us say some other place and to some other place and then reach the market. So, buses do not ply between any 2 locations.

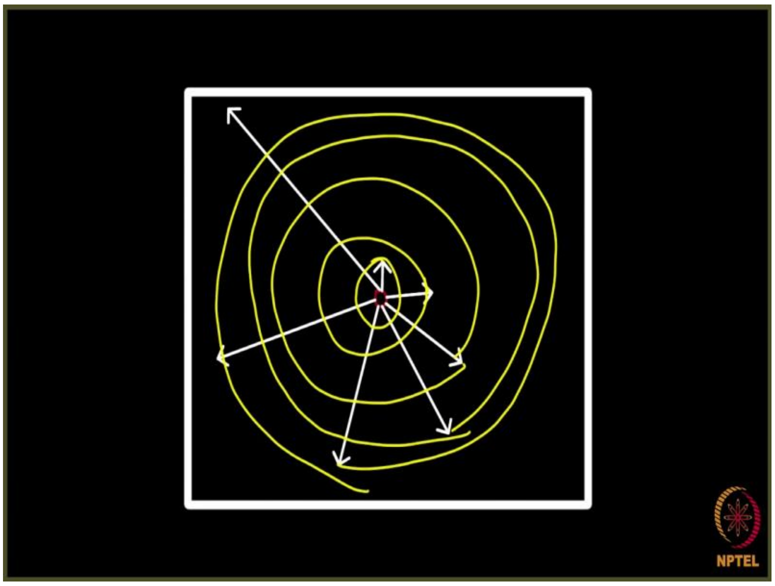
Now, here is the question if you were to decide what should be the bus routes, if you were to decide the bus routes what will be the locations from which you will um deploy a bus right. Assume every single location has the same amount of people and the requirement for them to travel is also the same, what exactly will you do? How will you plan the bus routes? Right, assume you have some 100 such locations here, there are 100 such red points whatever you are seeing here 100 such locations. So, whatever points you are seeing here assume there are 100 such points and you cannot simply put buses from any point to every other point right that is not feasible that is not practical.

(Refer Slide Time: 07:09)



So, what you do is this; assume this is the city all you do is you there is a central bus stand, let us say this is the central bus stand right. And, then from here you will have a bus go this place, this place, this place, this place, this place and this place right. These are the several locations where your buses go right and then, do you see something here? Do you see that? There are concentric circles here and you are trying to ensure that you have a bus for every possible concentric circle correct. So, what do I mean by this? By this I mean, if you let us let us now take a grid which is easy on the mind.

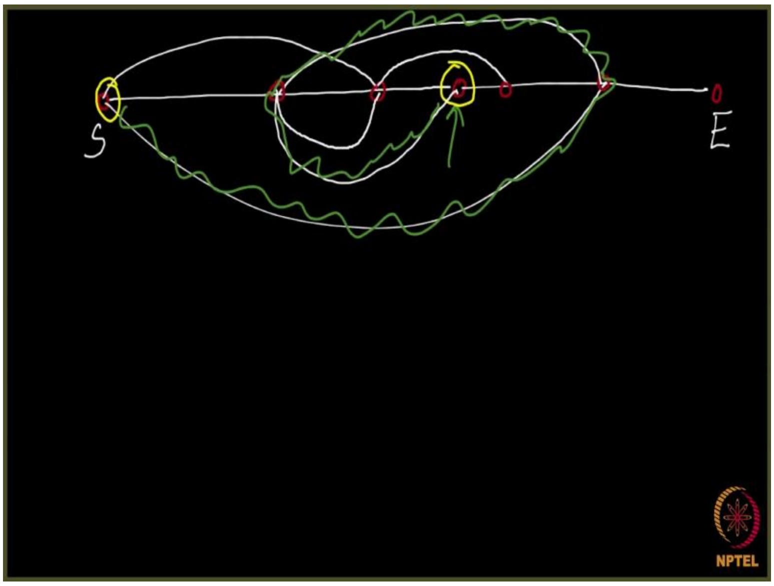
(Refer Slide Time: 08:01)



Assume there are some locations here, you would not sure that there is a um bus that goes to concentric circles to concentric circles. So, what do I mean by that? By that I mean from this centre let us say from the centre there are buses which go just to the next location, next to next location, next to next location, next to next location so on, so on and so forth ok. You will try to cover all possible locations from here correct, all possible locations. I am not drawing it well, but you basically you are getting the right intuition right.

So, that you are you are basically covering all possible areas in the place and the bus will go there and from there you may want to take another bus.

(Refer Slide Time: 09:03)

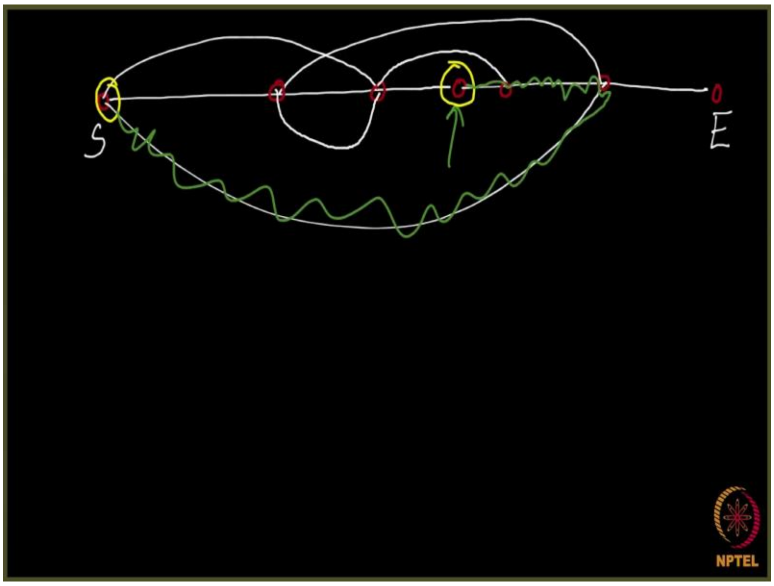


Simply very simply stated, assume there is a metro service where you need to go to let us say this place ok. Metro starts from this is the start and this is the end, let us say it is linear it is just it is a straight like this and you need to go here. You need to go here, but a metro train goes to this location and let us say it goes to this location ok, these 2 locations let us assume. But then you may want to get down here and then take another metro to another station, but that station goes a little further away and from here you can always come back to another station.

And that station probably takes you to further away place, there are probably isn't any metro from here to here. So, as you can see you need to reach this place starting from this place, but the metros are very weirdly distributed. Some superfast metros you know that they do not they do not stop they do not stop; from here you directly take you here. And, then from here you can only come here and from here maybe there is train to this place and there is not a train to this place. See it is complicated, how would you come then?

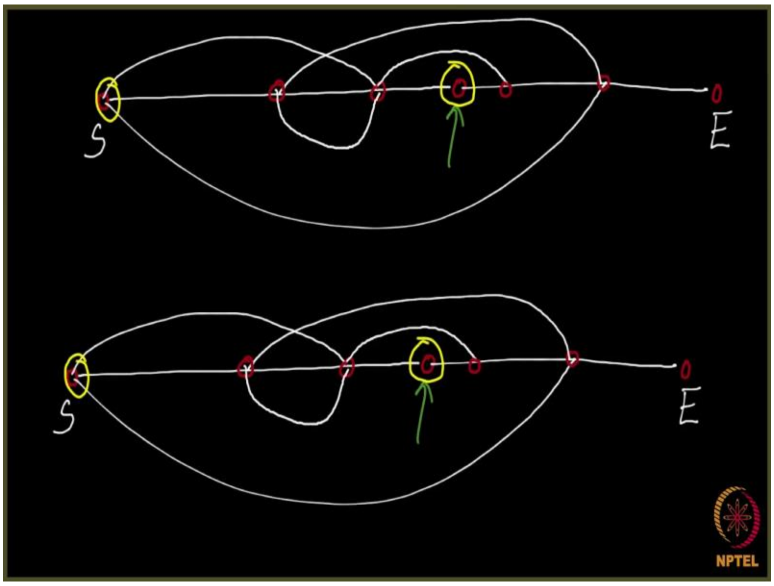
So, then you would probably take a metro to this place and then travel to this place and then travel to this is place. You know how to take a detour here; you see this is how things work.

(Refer Slide Time: 11:05)



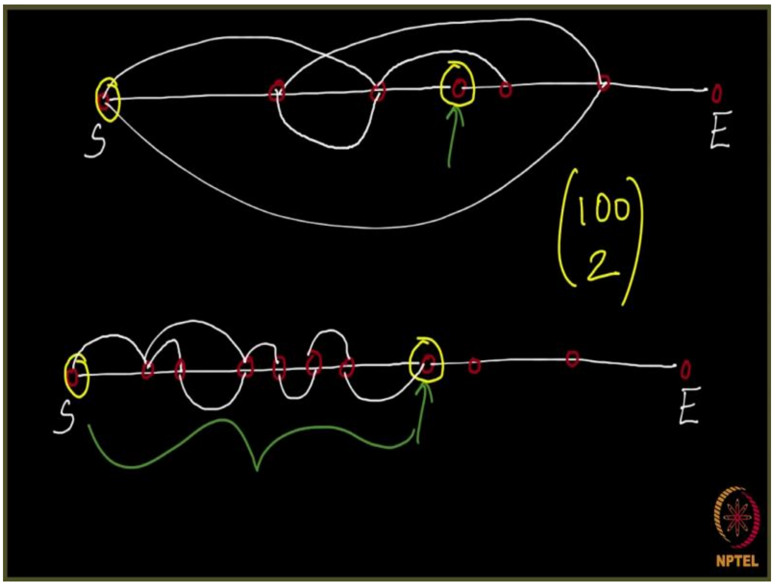
That is because if your reach is high, by that I mean see if you have long routes people may want to come back from at ok. If you only have shorter routes let us see what happens.

(Refer Slide Time: 11:15)



If there is a shorter route from S to E exactly the same thing ok, let me just that replicate this and then put it here so, that it is easy for us to work. I will copy paste this and let me give you another scenario where every single route is fairly short ok, it is fairly short. So, what we do in that case? That that comes with another baggage of problems right; let me once again copy paste this in this and then pasting and then I am going to tell you that, what if this were not the paths let me remove these paths now. And, I need to go these to same destination right, I need to come and reach here right. I am going to start here, but then the metros are all like this now, there is there is a metro to this place and then another metro to this place.

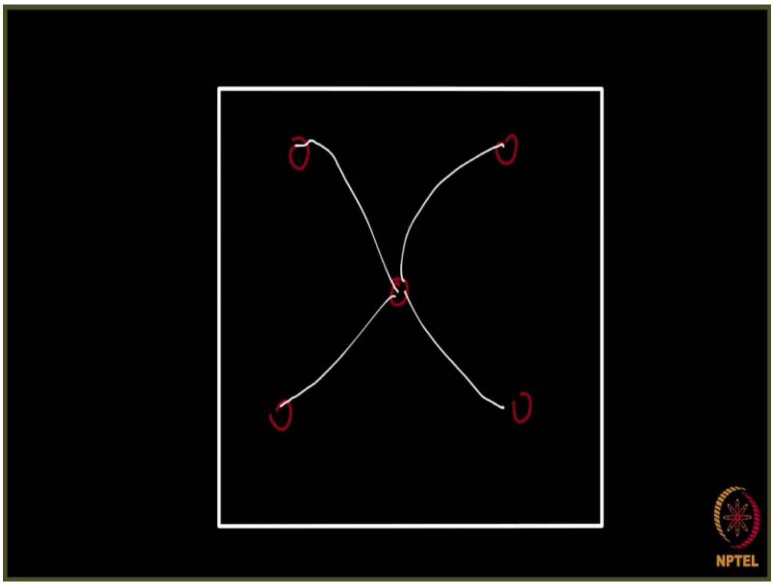
(Refer Slide Time: 12:15)



And then from here there is another metro this place so on and so forth. Here to here and then again from here to here alright. You know how to do a whole lot of toggling; you see a whole a hell lot of toggling you may have to do from here and then to here and then there is a metro to this place right. So, you may have to do a lot of jumps here, it this is a very strenuous process right. This is a very strenuous process, as you can see. Why is that, there is a no there is no quick way of coming from here to here right. If there are some 100 locations here you cannot put 100 choose to 100 c 2, let me write that down you cannot put 100 c 2 number of um trains or buses in this on this road. You can only put a few buses or trains between stations correct.

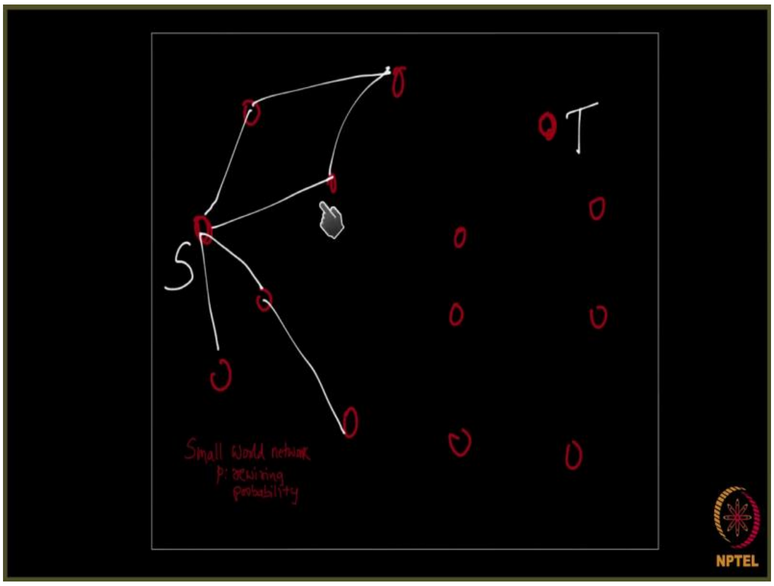
This the above one has a problem that, if you put superfast trains you may have to go ahead and then come back while, here you may have to do a lot of hops it is a lot of waste of time; you see you need to get in to a bus and then another bus and then another bus and so on, you need to do a lot of hops. What is the ideal way to construct um metro service here right um? So, what we will do is, if you observe this problem in this is the 1 dimensional, you see starting from here and going to here. If you see the same problem in 2 dimensions, if you see the same problem let us say assume you write a square or a circle you represent a city right.

(Refer Slide Time: 14:05)



We have exactly the same problem that when you want to put buses from one locality to the other right, it is a huge issue from where to where will you put from where to where will you put. So, let us do one thing assume you were to write program you accomplish this. So, what will you do? You will first generate a small world network.

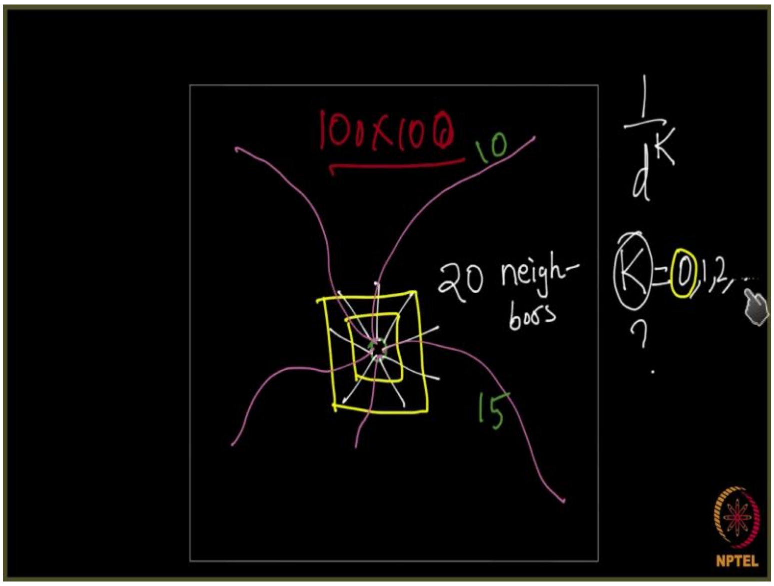
(Refer Slide Time: 14:55)



Let us say a small world network with some rewiring probability ok; p is the rewiring probability; you will indeed get some graph here right you will get some graph here right. So, let me let us write that down I get some graph here right, I get some graph here. Let us say a huge graph here correct and then I can go from any node to any node right node. But I must ensure that I can go from any node here to any other node here, from any source to any target I should be able to go from this source target that should be able to go. And, I want my edges to facilitate this that the whole point edges to facilitate this.

So, how does a small world network look like? It will have a lot of edges to its immediate neighbours geographically ok. Look at this is this is a city and a closer the vertices more are the possibility of edges, but then here and there I do put edges to further away nodes to ok. Now, what I will do is let me assume that I put I put um 10 I take 100 cross 100 grid right, 100 cross 100 grid whatever this was, if you remember a huge grid.

(Refer Slide Time: 16:33)



I will take here 100 cross 100 grid, then what I do is I um I basically will do this I will let me write that down; it is it is a slightly involved. You will understand it better if you get the next 2 minutes of my explanation. So, this is a 100 cross 100 grid let us say right and every single node here, let us say every single node here is adjacent to some let me say 10 nodes 10 nearest nodes; whatever they are on the grid alright 10 nearest neighbours.

What I mean by that? So, all these people are maybe at 2 levels ok, let us say some a 1 2 3 4 5 6 7 8, let us say some 20 neighbours 20 nearest neighbours ok.

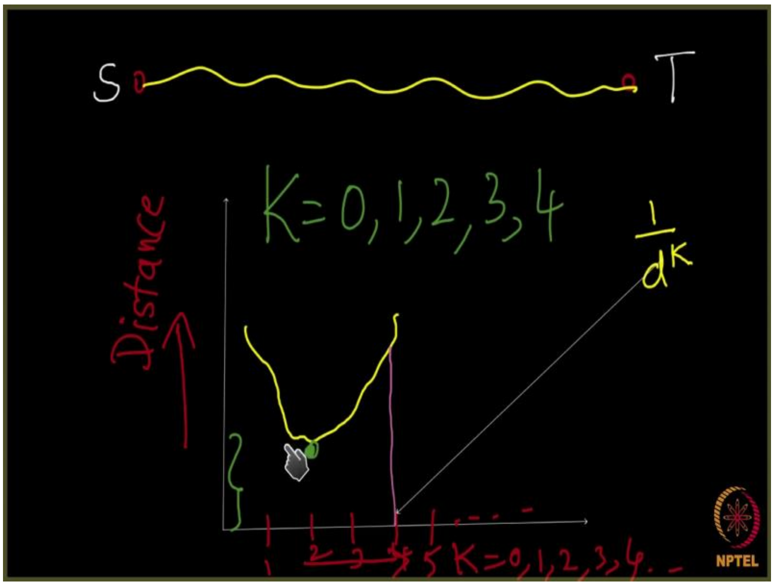
This is the homophile part 20 neighbours and then it also is adjacent to a few nodes further away and this is decided based on the probability function 1 over to the K, where K can be 0 1 2 etcetera as I told you right. What should be the ideal value of K? If I were to run this experiment on a 100 cross 100 grid right and I will choose first K equals 0, what will happen? When K is 0, this entire thing will be 1 which means it will be the same probability with which you will be picking vertices further away; its uniform distribution that can be dangerous you see.

Why? You will randomly put roads between 2 nodes then you may have to go longer distances and then come back into if you want to come to the destinations alright; it may not really serve your purpose. If you put K equals 1 then node that is 10 units away and the node that is 20 units away 10 units is more probable because, it will become 1 over 10 probability 20 units with will be a lot improbable, because it be 1/20 ok. Now, instead I take K equals 2 this be then it becomes let us say assume this was 10 units and 15 units so, 20 units whatever anything you want.

Then if you pick K is equal 2 this becomes, the probability of this being friends will be $1/10^2$, the probability of this being friends will be $1/15^2$. But, then again there is a problem there would not be very further away nodes because, $1/15^2$ is 1/225 and that is a very small number, which means a node that is 15 units away by units as you all know, I mean the number of edges the shortest path length. If a node is 15 distance away then the probability of that being friends is $1/15^2$, but a node that is 10 units away is $1/10^2$ which is 1/100.

This is more probable; this is very less probable than what it is used to be when K equals to 1. As you can say if K is equals to 3 or more nodes that are far away becomes very in probable and navigation becomes difficult for you to go from a source to a target. So, here goes our question we were talking about simulation right continuing our discussion.

(Refer Slide Time: 20:43)



If you were to go from a given source to a target T target node T and you are considered different K. K is K equals to 0, K equals 1, K equals 2, 3, 4 and assume you draw this plot ok. Assume you draw this plot of this y axis versus the x axis right and then you will plot, what are the path lengths. If you put on x axis you put K which is basically 0 1 2 3 4 etcetera ok, 1 2 3 4 5 and so on. And, on y axis you put the let us say the distance, the distance covered when you take the check the typical such algorithm. When you when you go to the next node and then you check how far is the target and you go to the next node that is closest to the target and so on.

When you do this, you will observe that when K equals 0, K equals 1, K equals 2 etcetera the plot looks like this ok. It let me write nicely, the plot has the observation is at 2 it says something. The distance is minimum now what, but do I mean by distance? By distance I mean you need to the destination you can go and reach quickly. If the rewiring probability is $1/d^K$, where K is 2 when K is let us say 3 the distance is slightly more as you can see. When K is let us say 4, then the distance is a lot more as you can see ok. I hope you are understanding what I mean by this distance.

By this distance I mean the path length of the process, when you want to go from the source to the target it is a longer path if you were to choose K as 4. If you were to choose K as 2 the shortest the path from S to T, what do you mean by what do I mean by path from S to T? The same way Milgram experiment, when you take a node and you go to the next node which is closest to be on the grid ok; of all the neighbours you choose that node that is closest to T on the grid. That is what we have been discussing all this while right. And this particular algorithm will tell you, this particular programming exercise will tell you that a K equals 2 is the best possible value for K.

Whenever, the rewiring probability that you are considering is given as 1 over d to the K right. So, if this is sort of a hand waving that I have done here. In fact, that is a very beautiful proof for this saying that K equals 2 for grid and K equals 1 for 1 dimension. And textbook describes it, it is very nicely described with some very basic probability you can understand it completely. Let me know if you have not understood yet, I will help you out understanding it. But, this will be out of your own interest; this will not be part of the course examination.

But, the idea here is simple idea here is that you use K to sort of decide on whether you will make friends with nodes that are further away right. And, when you make friends with nodes that further away a lot more then you are in trouble; a lot less you are in trouble. The perfect balance is when you choose K equals 2 ok. In fact, we will be showing you programming exercise; we will be explaining this in detail. It will become probably clear to you, then. As of now just know that this is curve is how it looks like and for 2 the distance is the shortest.