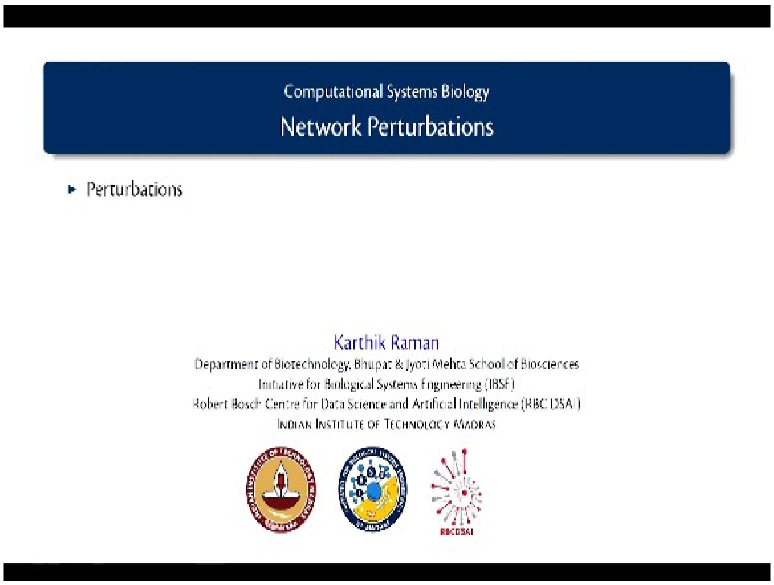
Computational Systems Biology
Karthik Raman
Department of Biotechnology
Indian Institute of Technology - Madras

Lecture – 22
Network Perturbations

(Refer Slide Time: 00:11)

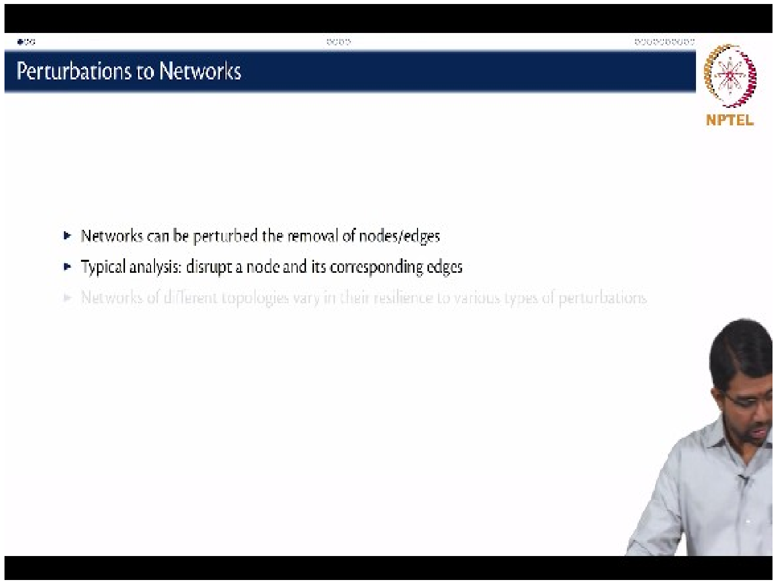


In today's video, we will start studying network perturbations. What are the different kinds of perturbations that you can do to a network and how do you measure the response of the network to these perturbations? So, first we will look at perturbations. How do you perturb a network? As we discussed earlier, perturbation is one of the most important themes in systems biology, in biology in general, right.

You try to knock out a gene and see what happens, you knock out a bunch of proteins and see what happens and so on. Similarly, here we would like to see what happens when you perturb a particular biological network. How would you perturb a network? So, the most classic way to perturb a biological network would be to remove a node alongside all its edges.

You select a node, then remove it which means all its interactions also get lost. The other interesting way to perturb a network would be to just target a single edge, remove just one edge out of a network. It is a very selective perturbation while the other idea is more of a wholesome perturbation.

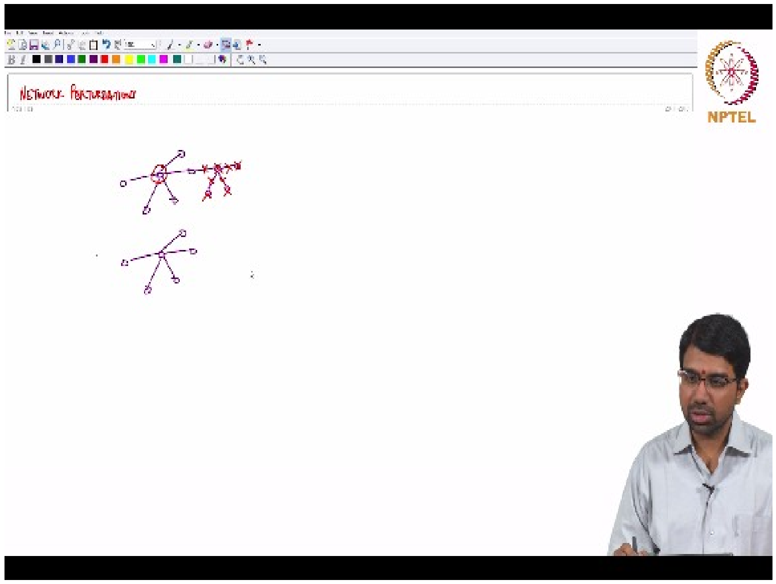
(Refer Slide Time: 01:14)



So, we can perturb a network by removing nodes or edges or disrupt a node and all its corresponding edges but networks of different types will basically behave differently when you start perturbing it. **"Professor - student conversation**" You can add networks but in general, add nodes and so on but in general, we study the perturbation to removal.

So in biological networks, well, you are often interested in trying to find out what happens when something fails, right? You can also study what happens when you add something but more than adding something may be in a biological network context, one interesting perturbation would be what happens when a gene duplicates, okay?

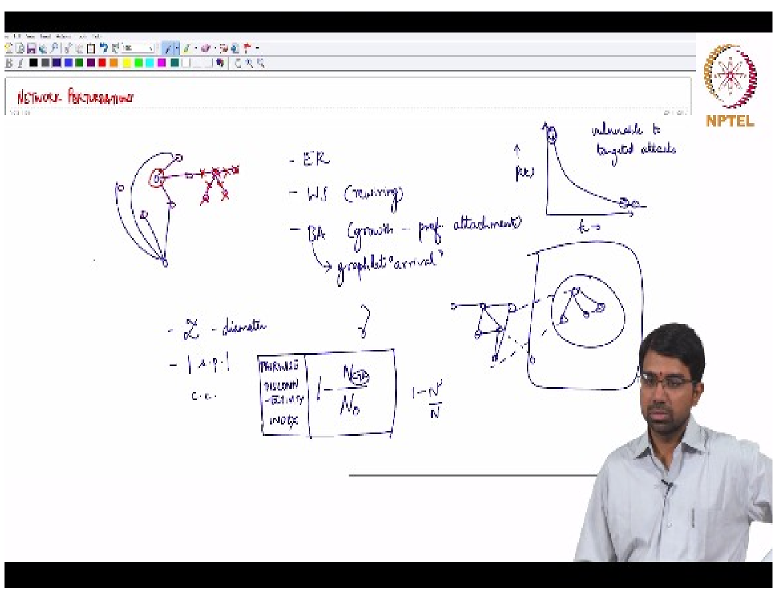
(Refer Slide Time: 01:59)



So, what do you think happens when a gene duplicates. So, let's look at network perturbations. So, you have a network. It looks like this, right. So, obviously when you perturb the network, let's say you remove this node, you essentially remove all of these, right. So, these nodes will also go out therefore because they don't have any other connections, right. But what happens when you duplicate something, when you duplicate an edge, right or duplicate a node.

So, let's say you duplicate this node. What will happen is something quite interesting. You have now duplicated that node, right. Is there anything else that should happen? **"Professor - student conversation"** It will basically interact with most of the previous partners, right.

(Refer Slide Time: 03:14)



So, it's not going to be interacting with these nodes but in fact, it is going to interact with, so you basically have now duplicated that node but it is going to interact with all the previous nodes, right, because it is just a duplicate. It is very likely to interact with that. But then in future, other interesting things can happen. May be there is no selection pressure on this edge, so it can go away.

May be this edge could go away, this edge could go away, right and your resulting network could be very different from what you started off with, right. So, in fact you can also, so we studied a few ways of generating networks, right. So we studied ER networks. Then we studied WS networks which was, which had some rewiring. Then we studied Barabasi-Albert networks which had growth by preferential attachment, okay.

There are other models that we haven't studied. So even in Barabasi, the classic example that we sort of looked at was, what happens when you add 1 node at a time? To start with an initial network and keep adding one node at a time but you may have what is known as a graphlet arrival. What is a graphlet? It's a subgraph, right. So, typically when you look at social networks, that becomes very important.

Not so much in biological networks, except if you have community networks and so on, right. So and obviously social networks are also of interest in biology if you want to study the transmission of a disease and so on. So, this basically means that you have some existing network and this is the network that arrives. It is not a single edge.

So, we said that one edge arrives and it will connect with probability proportional to degree and so on. Instead you have this graphlet or this subgraph arriving, right. So now this might connect in different ways and so on, right. So, this could be potentially another network growth model and this is, this is indeed a sort of perturbation to the network in relation to what you just asked.

"Professor - student conversation" You can have more edges. So, that will change the γ. If you keep adding one edge at a time, I think you will get γ=2. You can just plot it and check it out. **"Professor - student conversation**" Yes. Yeah, yeah, so that is why I mentioned this. So, this is essentially a perturbation, adding this in network to the original network is a perturbation.

So, you can perturb by loss of nodes which is the most common way to perturb the network but this is potentially a perturbation that you want to understand as well. So, how do you quantify network perturbations? So we have a network perturbation, how do we go about quantifying the perturbation? You can study most of the properties that we have studied so far, right.

But you might want to start with say characteristic path length, diameter, another useful metric is number of shortest paths. We can of course, look at clustering coefficient, centrality measures, so on and so forth, right. You can study perturbations to a network in many different ways, right but the most useful, most commonly used methods are these. So can you study diameter, characteristic path length and shortest, number of shortest paths.

So, now go back to the networks that we have studied so far. We have studied 3 types of networks, right. A random network, a Watts-Strogatz small-world network and a Barabasi-Albert power-law network. Which of these networks do you think will be very susceptible to perturbation or will be most resilient to the loss of nodes or lengths and so on?

So, Barabasi-Albert will be most robust. Why? Robust to what? You have to qualify your statement. To random perturbations, right. You randomly remove a node from a Barabasi-Albert network. It is not going to affect it. Why?

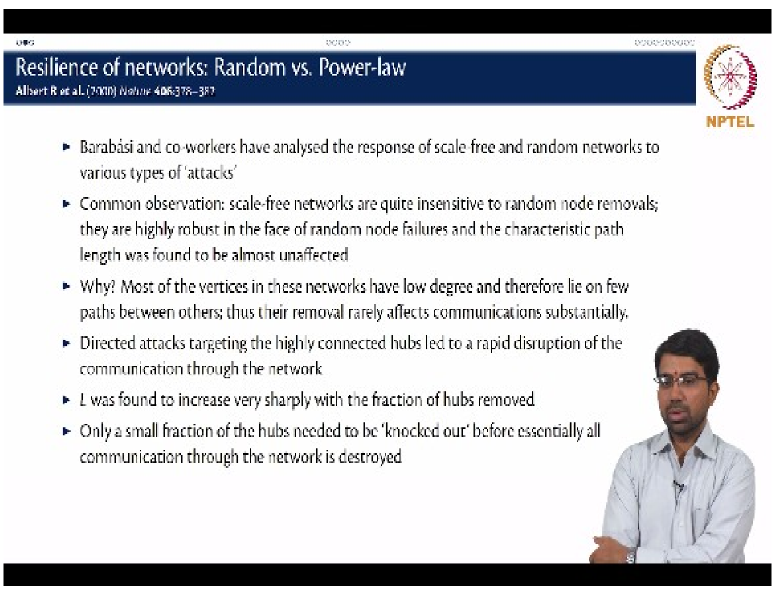
Because a random node is likely to have low degree. Because the degree distribution looks like this, right. So, these are the most populous nodes, right. So, once again this is k versus p(k). So, these are the most populous nodes. So, you are very likely to pick one of these nodes, pick one of these nodes when we hit at random. The problem for the Barabasi-Albert graph comes when you start hitting these nodes.

So, Barabasi-Albert networks are vulnerable to targeted attacks. You need only to knockout a few hubs before you disrupt the entire network. Whereas in a random network, there is not much connectivity to begin with. So, when you start knocking out, you really don't, it slowly starts disintegrating. The characteristic path length will slowly increase. Increase or decrease?

It will increase. Diameter will again? Increase. So that is a fair point. So what will happen is? The characteristic path length will slowly increase and then start decreasing. So, initially when you start knocking out a few nodes, what will happen is your components haven't yet broken down. You still have like some biggish components and those components are getting poorer and poorer in connections.

After that, it becomes islands, right and then it kind of decreases. So, you will have like an increase and then phase change sort of, similar behaviour, identical behaviour. So, diameter will initially start increasing. At a point when you know you almost have like very long paths in a very sparsely connected component, then it will come crashing down when you, when the components just fragment.

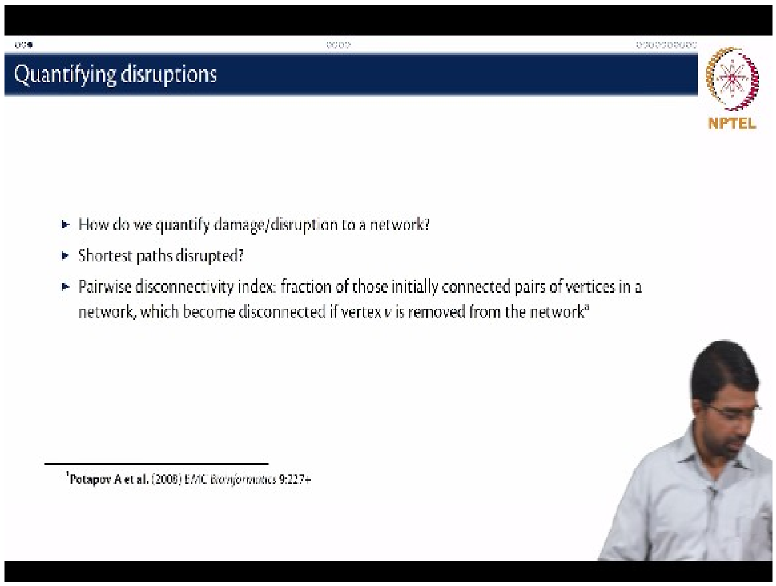
(Refer Slide Time: 10:38)



So, Barabasi and co-workers studied scale-free networks and random networks in terms of how do they respond to attacks, various types of attacks. So, you could have a targeted attack as we discussed or you can have a random attack. Common observation: scale-free networks are insensitive to random attacks. It's not surprising. Because most vertices in these networks have low degree. Whereas directed attacks targeting the highly connected hubs will rapidly disintegrate the network. So, these are experiments you can try, right.

You can run a simulation. You can create a scale-free network, start knocking out nodes at random and then say plot characteristic path length. So, L was found to increase very sharply with the fraction of hubs removed, right. As you remove the hubs, it presents, it increases very fast and then you have a drop and the whole network gets fragmented and so on and you needed only to knock out a small fraction of the hubs before literally killing the entire network.

(Refer Slide Time: 11:40)

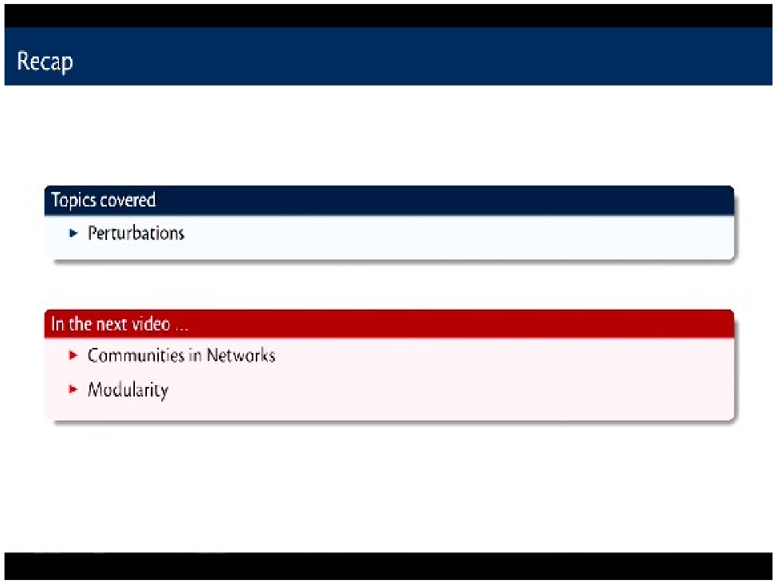


So, how do we quantify the disruption? So, one way of quantification would just be computing a delta characteristic path length, delta diameter or something like. You could also look at the number of shortest paths disrupted. There was another method that was proposed called pairwise disconnectivity index. This is basically nothing but let's say this is the number of initial shortest paths in the network, right, N_0 and you now removed the node v, right.

Now, how many shortest paths remained in the network? So, this is the metric called pairwise disconnectivity index. You can come up with your own metric for trying to understand how a network disintegrates on the removal of nodes. Yeah, so whatever. You can just have some N or let's just say 1-N'/N. N is the number of shortest paths in the network. So, if you have your all shortest paths matrix, right. So, this looks at pairwise shortest paths, right.

So, this somehow measures a pairwise disconnectivity. Some pairs were initially connected but now they have become disconnected or some pairs had many more paths initially, now they have much fewer paths. So, you can measure it in 2 different ways. So, the classic way is to measure how many nodes gets disconnected but you can also count how many shortest paths get lost. Maybe you have alternate shortest paths which gets lost and things like that.

(Refer Slide Time: 14:05)



In today's video, I hope you had some insights into how one tries to perturb and measure the effect of perturbations on any given network and more so a biological network and in the next video, we will start looking at communities in networks and also modularity.