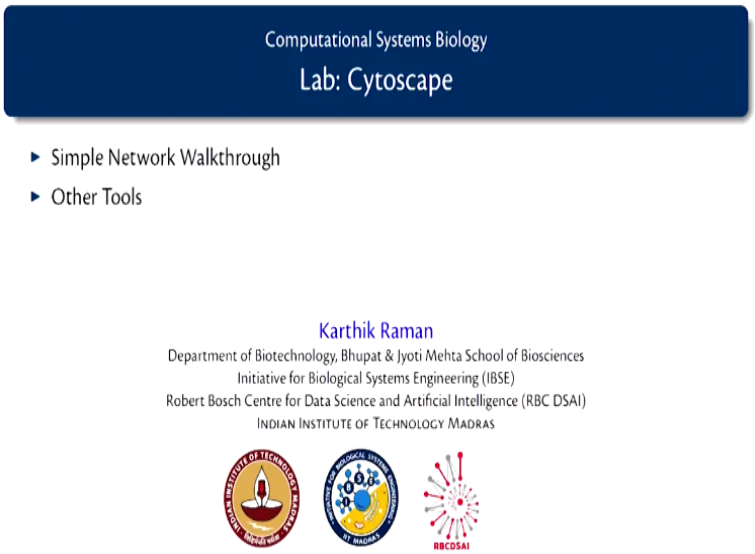
Computational Systems Biology
Karthik Raman
Department of Biotechnology
Indian Institute of Technology – Madras

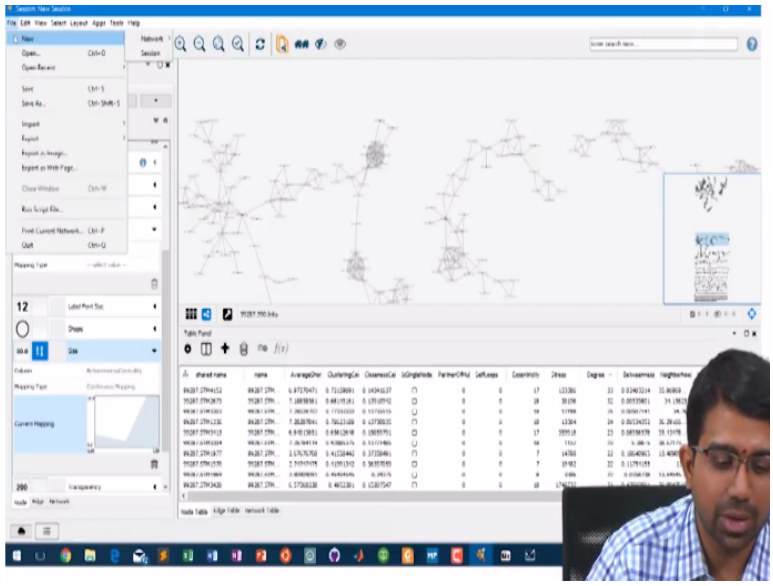
Lecture - 26
Lab: Cytoscape

(Refer Slide Time: 00:11)



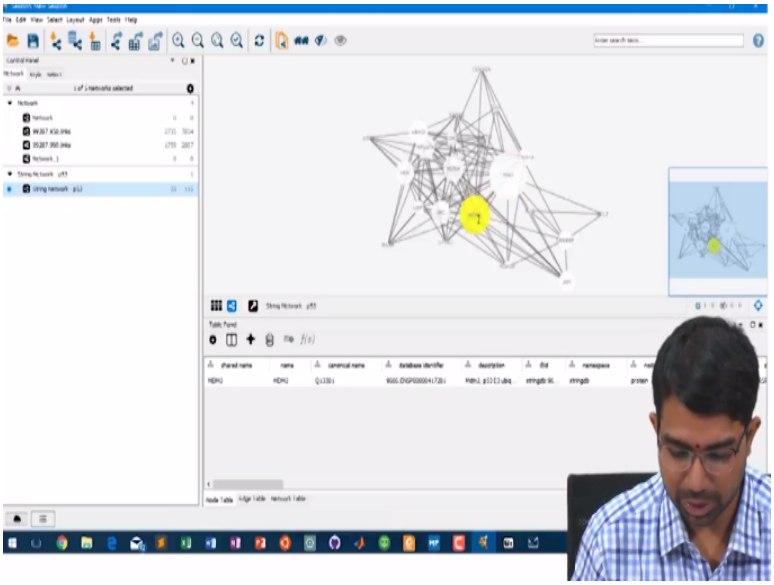
In the continuation of the Cytoscape lab, we will have a walkthrough of analysis that you can do with a simple network and I have also briefly introduced you to other tools like Gephi and so on that exists for network analysis. There is also a very interesting MATLAB tool box called MATLAB BJL which we will look at in a later class.

(Refer Slide Time: 00:34)



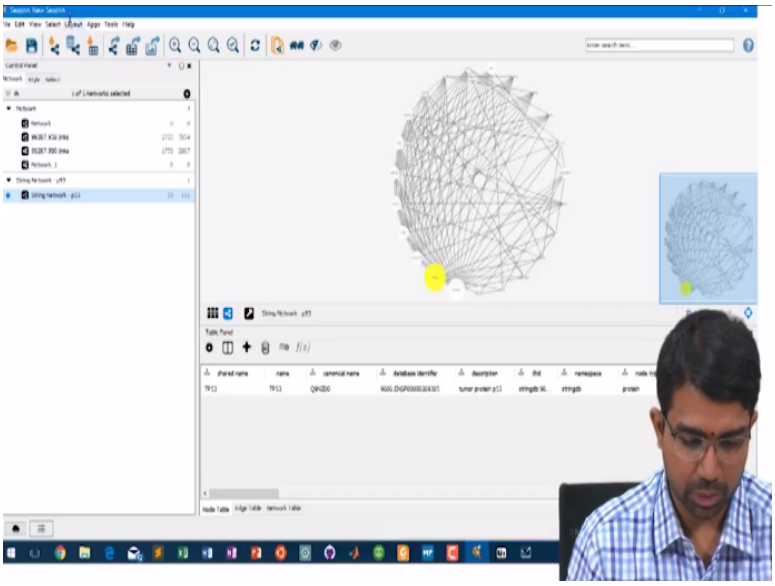
So let us may be play around with a simpler very simple network.

(Refer Slide Time: 01:28)



So now I know what node is the highest between a centrality and so on, like I can have different types of layout algorithm as well. **"Professor - student conversation starts"** So we can fit in our algorithm? For a layout, no. There are built in algorithm you can write an app if you want. You can have all kinds of layouts that you can usual nice layout is organic. It spreads out usually quite nicely. **"Professor - student conversation ends"**

(Refer Slide Time: 02:09)



There are various layouts that are possible. You can manually lay it out which you may often do for a smallish network if you are trying to explain something or you know preparing an illustration. This, as you see are substantially large network that is about 3900 nodes and you can easily analyze this also using cytoscape. So you see these are pretty fast.

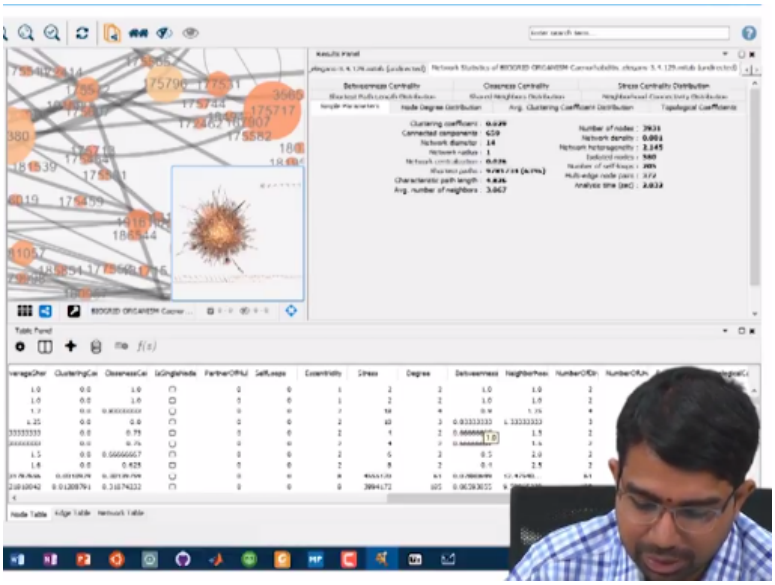
(Refer Slide Time: 03:40)

Let us look at the degree distribution and let us try it power law. It has got a good correlation and r squared well. When you look at the degree you have points even with a degree of 524 huge the one know which has 524. Let us try to visualize that you can easily capture where that node is? So, map node size to degree and map node colour to between a centrality. If you want, we can even map edge size to edge between this. There you go. So let us know zoom into it.

It is going to be little tricky. It is really a dense network. So what we can do here is you can just select nodes and you can look at all the so for example let us say you want to let us again do the network analysis, network analyzer, analyze network, undirected. Let us look at some parameter like characteristics so what is the network diameter is 13. So now let us take the top degree nodes.

So we have something that is 524, 1, 2, 3, 4, 5, 6, 7, 8 9, 10. I will now pick the top 10 degree nodes so I am only selecting these nodes and delete those nodes. So let us thus probably just take a screen shot of this and let us again analyze the network. Let us just look at least the diameter was 13 and characteristic path in the so 0.321. Let us see how that changes.

(Refer Slide Time: 06:50)



So for 4.321 it became 4.826. So the diameter just varied by 1. Not a big deal. Now let us take between the centrality. **"Professor - student conversation starts"** Again we have to analyze from the beginning right? True, but we can just, we are doing some random perturbations here so we can try and mix the pertubations. I think this is between a centrality of 1 let us see what these nodes are I am not sure of their, they could be nodes in this other components. May be we should restrict ourselves to this component.

“Professor - student conversation starts” It is possible to take the single most character component. Big single most character component. So there is one way to do it so let us take this node. I know it is a part of that connected component and I just go to edit so tools and select and nodes first neighbours of selected nodes. So I can do control 6 so that has now selected this entire component. Control 6 was to select the neighbours of my nodes. **“Professor - student conversation ends”**

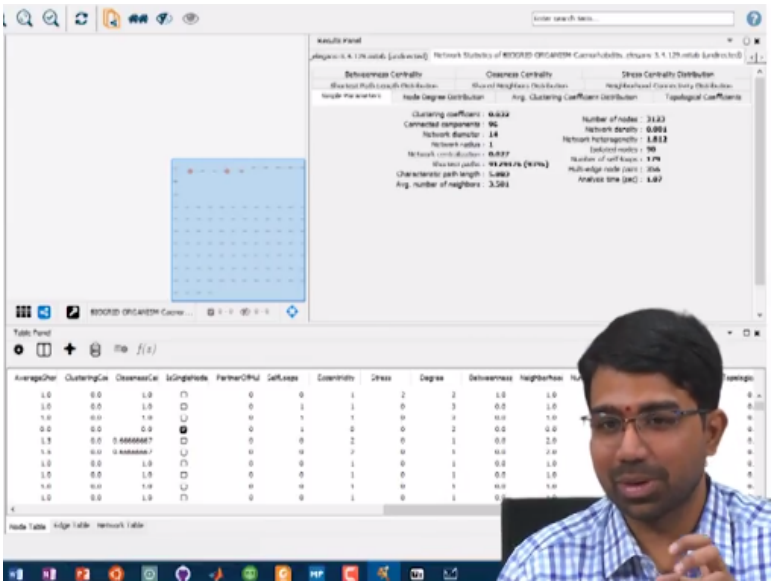
I did that recursively you do that a few times you did that 14 times basically that guaranteed to capture the entire network so now edit select nodes inward selection control I. Now I have just the single component. So now let us look a shortest path it is 662% now I think that should come up to 100%. Tools network analysis, analyze network. So, all nodes are obviously connected by connected to one another.

“Professor - student conversation starts” Parameter is the same. Diameter will decrease? No, Diameter won’t decrease. Well not necessarily, so if you had another you know cluster which was like really bad linear chain you could have the diameter decrease. Let us now look at the degree reservation. **“Professor - student conversation ends”** So clustering coefficient network clustering coefficient is 0.037, but let us look at the CK versus K. This is interesting.

Now let us try to now let us take between S the highest between is only this much let us take let us see if this even breaks it breaks the network into some more components. Simple parameters and tools. So now we already had some disconnections. This is the shortest paths has gone down to 93% by just picking the top 5 between a centrality nodes I am deleting them I think we created some more components of the network.

“Professor - student conversation starts” But the diameters are changing. Diameters is not changed not necessary that it should change. Those components are smaller. **“Professor - student conversation ends”** At least the path length should increase right? So let’s just try to delete this and we will see. Yes, we have so many disconnected nodes. I delete in the single largest component and now you have so many disconnected nodes.

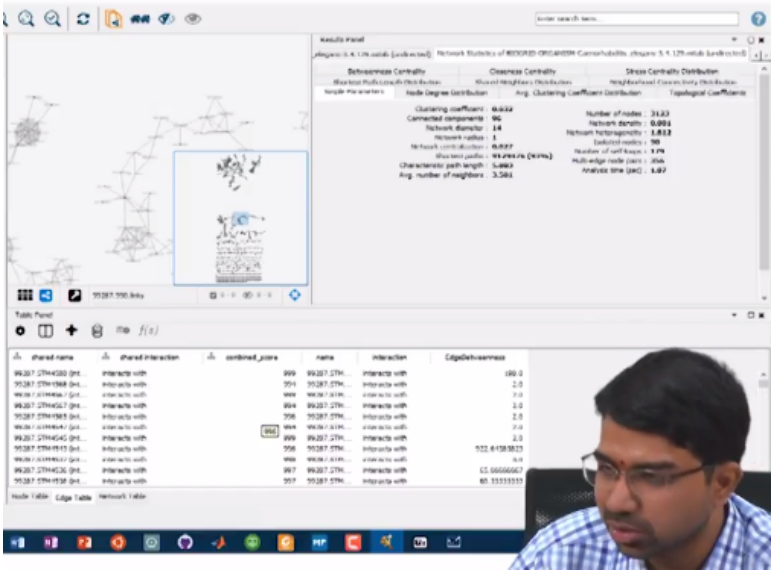
(Refer Slide Time: 11:43)



Let us lay them out again. It is completely disconnected because of you probably you know they did not have any other neighbours other than that sort of hub nodes to which they were connected, but essentially if the idea through this sort of lab session is to highlight what are the nice things that one can do is cytoscape so as you must have understood today it is a very powerful tool for visualization that you can easily visualize any parameter that you want.

You want to visualize between a centrality or closeness centrality or degree it is just 1 click, but you know how to do it for more complex things. The other useful thing to note is you have an edge table and a network table here and a node table. So all the node properties will be here any edge properties will be here usually not many edge properties I think you will have edge between this.

(Refer Slide Time: 12:36)



These are basic things may be let us look at a simpler network like the one we generated. So there are actually no edge properties at all that is like interacts with combined score. The score is of course an edge property and so if you want even here you do things like, you can select all the nodes with the highest score and so this was a 990 network. You can already just you are going to keep only the 999 edges.

One of the best way to do is it as you can imagine, what you get select it and deleted it from here and so on. So I have gotten rid of all those edges. So even useful for smaller networks if you are analyzing a particular network these some useful analysis you can do with cytoscape are like what you can do with the string app. You pick a particular protein of interest SAP53 or whatever.

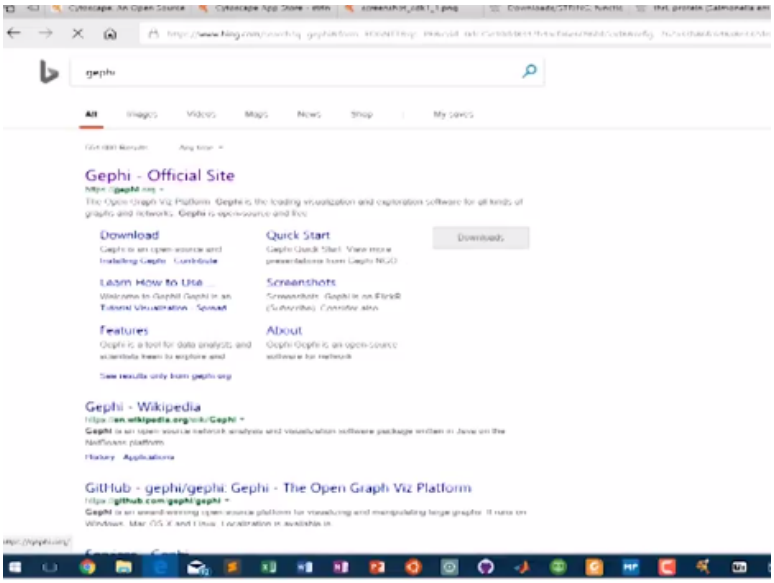
And look at the immediate neighbourhood around that because there are lot of interesting action happening there and that you can sort of analyze very nicely using cytoscape. Look at the immediate neighbourhood and you figure out what are the important proteins there and so on. One hop or even 2 hop neighbourhood. I think that is quite useful to look at. So hope you have understood this visualization part.

That is very useful how do you do a discrete mapping or pass through mapping and continuous mapping and so on because this is central to visualizing any complex network. Usually have some issues with this zoom in, zoom out and so on. Occasionally that is an issue but otherwise I

think cytoscape is very powerful. **"Professor - student conversation starts"** I think we can do discrete and continuous mapping in MATLAB BJL. How will we do a passthrough mapping?

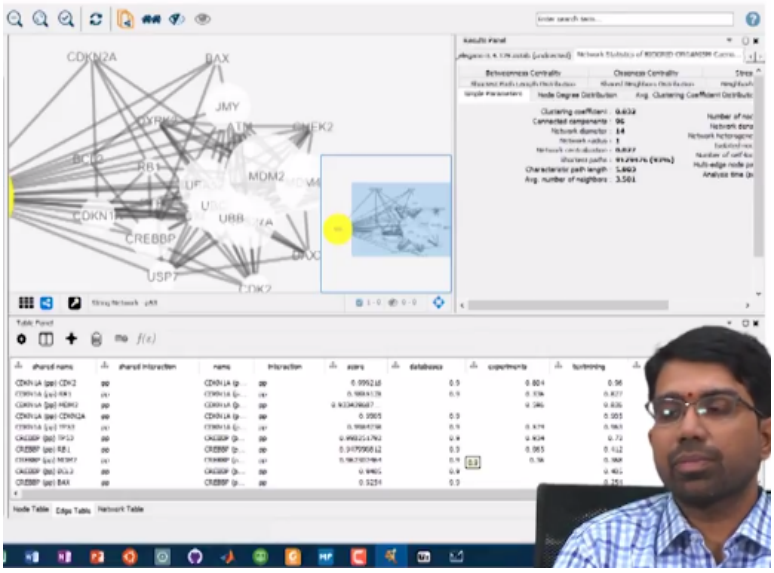
How will you do discrete mapping and continuous mapping in MATLAB? you cannot plotting is very poor in MATLAB. How do you even visualize it you do not have any nice visualization algorithms may be there are a few bundle it is not a part of BJL but it may be part of the MATLAB but NetworkX is good. NetworkX has several visualization algorithms. **"Professor - student conversation ends"**

(Refer Slide Time: 15:02)



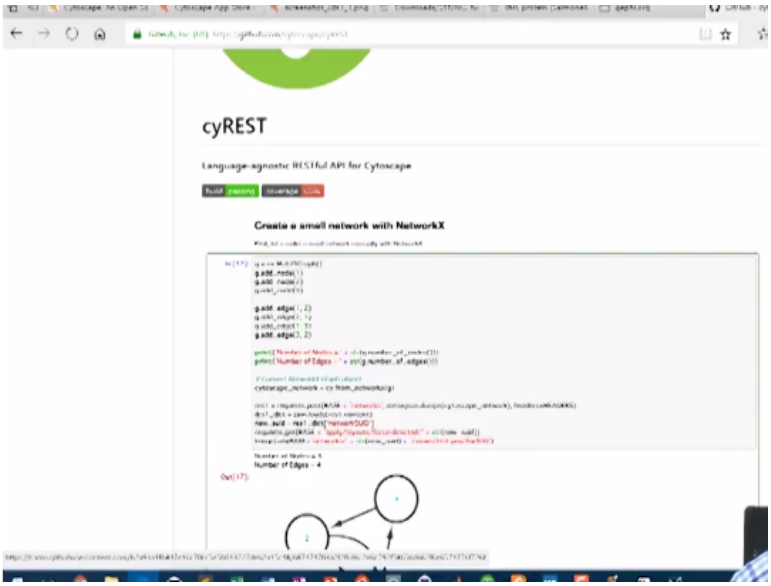
The other tool that I think what I want to study is Gephi. So very powerful tool for doing all kinds of network analysis and again you can make pretty nice pictures with it and so on and so there are many tools I think the most popular amongst them are Gephi and cytoscape. The good thing is that they are all cross platform random. You can run them on windows, MAC, LINUX whatever you want. There are many clustering algorithms that you can run as well.

(Refer Slide Time: 16:47)



There are also a few things you can do in a programmatic way as well as cytoscape so that is quite useful as well. You can access several of this cytoscape commands programmatically. I think there is also a terminal and so on. So there is a command end dialog I can do several things. So I think cyREST tool is also useful.

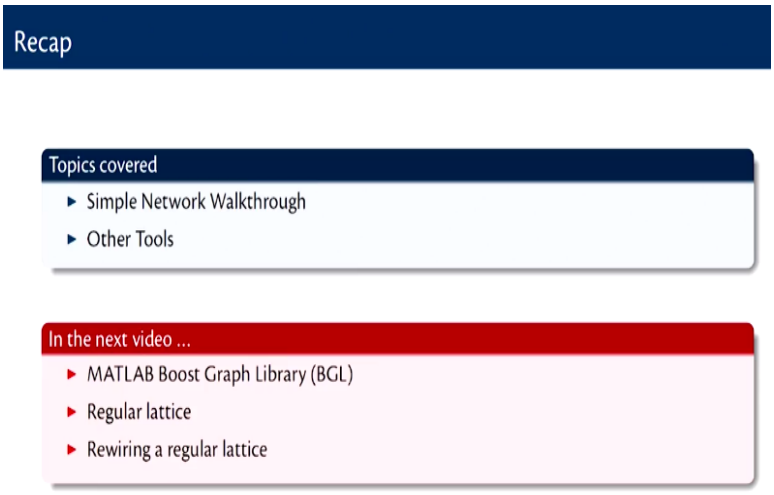
(Refer Slide Time: 17:33)



This also talks to NetworkX and after that you can use cytoscape tool. You can apply layouts network basically the GUI commands. I think we will stop here for today and this was basically to recap in order to follow up with the networks module and this is something that is important to know what are the all tools? So you should explore Gephi and cytoscape and any other tools on

your own. So cytoscape you had some sort of an introduction today, but there are other useful tools that exist.

(Refer Slide Time: 18:32)



So I hope you have a good understanding of how do you work with at least simple networks and cytoscape and the concepts are not very different if you want to work with a more complex network although it will get you know little less responsive and you will have more challenges and visualization and so on and there are many more tools other than cytoscape that one can use to study networks.

The next video I will introduce you to we will continue with another lab where in we will look at the MATLAB boost graph library which is an old very useful tool box for MATLAB to analyze networks. We will also look at how do you build the regular lattice in MATLAB and how do you rewire it and so on.