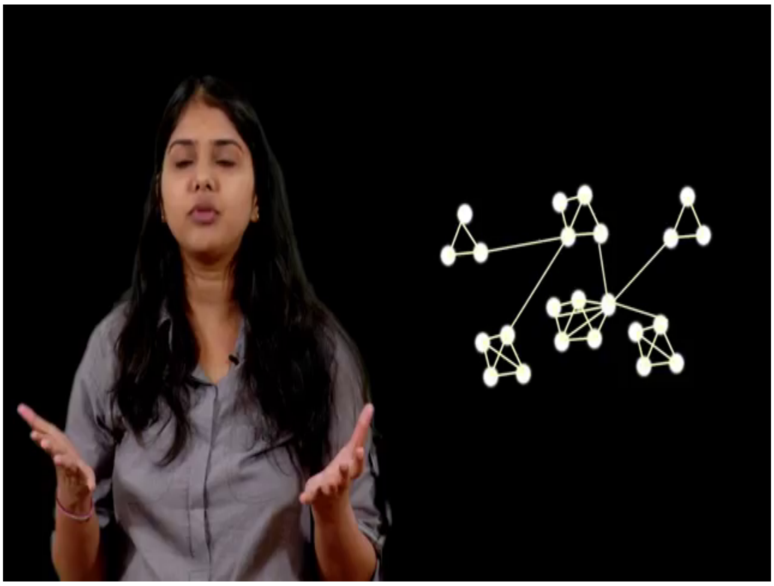
Social Networks
Prof. S. R. S. Iyengar
Prof. Anamika Chhabra
Department of Computer Science
Indian Institute Technology, Ropar

Cascading Behavior in Networks
Lecture – 93
Cascade and Clusters

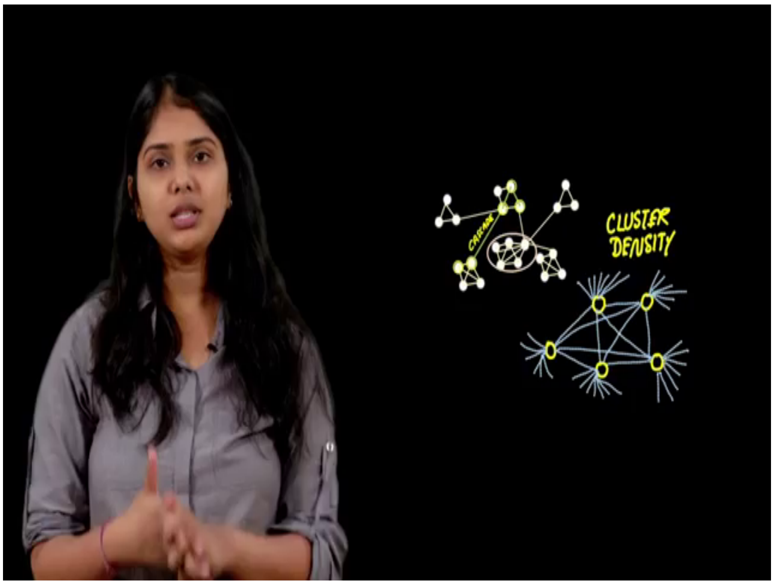
So, we have just looked at a story and have very easily said that, if there is a community having a high density probably and piece of idea traveling through this network cannot enter this community. We intuitively know this, but now let us try to look at it mathematically and it is actually very easy. So, to prove it mathematically, let us first refine a scenario and revise our definitions. Let us refine our scenario. So, our scenario was we had this class right; previously we looked at just 2 communities.

(Refer Slide Time: 00:42)



We said that this class where people are deciding whether to be in library or whether to be outside and then we looked at just 2 communities. What we do now is, let us take let us make this a little bit more detailed. Instead of just having two communities let us have a lot of small, small clusters in this entire network.

(Refer Slide Time: 01:01)



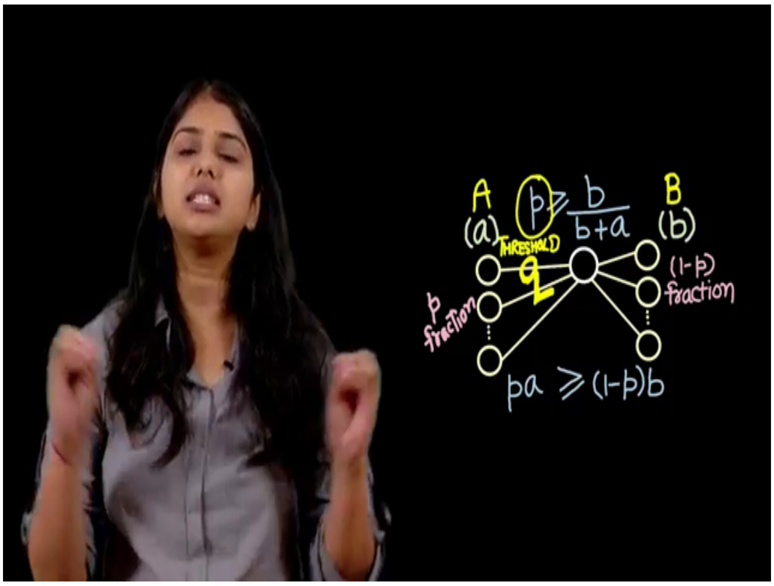
So, here is this class and people have made small, small clusters, it is actually what generally happens. A class is generally not split in 2 halves, people have their own bunch of friends and if it is a class of 100 people everybody has some bunch of 6 or 7 friends and people exist in small, small clusters.

And now the question which we ask is, in this class there are these 2 people somewhere in some part of this network who decide to go outside and have fun. And the question which we are asking is, will this idea of going outside this decision of going outside and having fun quickly sweep the entire class or will here be a bunch of people 1 cluster who despite whatever will not be able to convince to go outside? So, despite all of the rest class despite the complete of the rest class is going outside, these people stay together and say that we are not going outside and hence this diffusion process. So, we call it a cascade.

Now, whenever an idea or information diffuses, this the trajectory of this diffusion we get we call it a cascade as shown in the figure. So, we look at whether it becomes a complete cascade or not. Simply whether everybody in the class will decide to go outside and have fun or this bunch of people will not be convinced. So, this actually used to happen in our B. Tech class where it was not about sitting in library or going outside, it was about whether to have a bunk or not.

So, whether to have a bunk or not and then there were 3 girls in the class whose friendship was really, really strong and they would always say that no we are not going to go for bunk and the entire class had to cancel their decision. So, keeping this example apart, just for fun our question was whether this will be a complete cascade or not. So, for looking at it let us revise our definitions.

(Refer Slide Time: 03:18)



So, we have previously seen that if we look at 1 node, 1 person in this class, this person decides whether to go to library or whether to go outside based on the payoff functions and the number of friends. So, we have there come up with a value called threshold. And we have looked at if the fraction of the, fraction of these persons friends who are going outside exceeds this threshold, this person will also go outside.

So, it was for so it is 40 percent let us say, so it means that if 40 percent or more of my friends are going outside and having fun I also go outside and have fun. And if 60 percent of my friends have to, are sitting in library and working then I sit in library and work. So, the threshold is 40 percent for going outside and 60 percent for staying in the library and working. So, this value threshold we denote it by a number let us say q. So, q is a threshold associated with every node. So, at least q fraction of this nodes friends should be convinced, should be adopting a particular decision for this node to adopt the same decision.

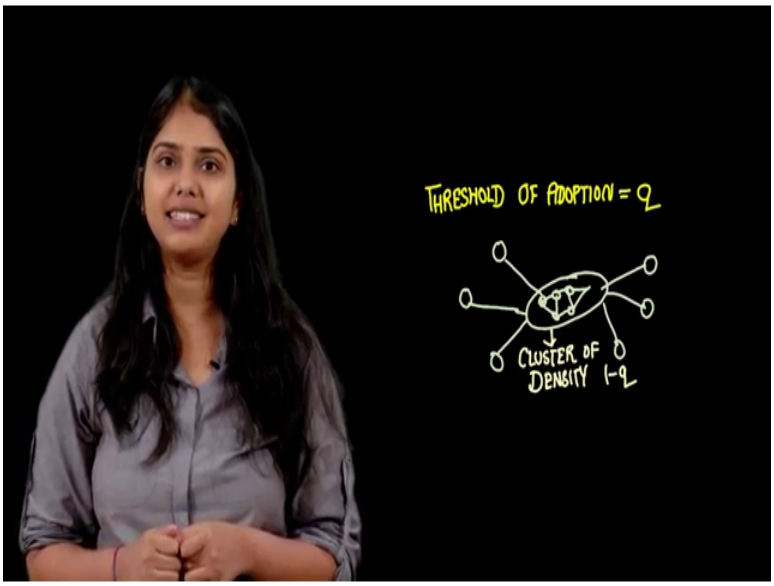
Now, what was our question? It was related to the cluster. So, we have these clusters in this class and in every cluster there are some nodes and each node is having some threshold value q. So, here we are assuming that every node is having the same threshold value q, I will come in on it later. So, every node in every cluster is having this threshold value q. Now the next important thing as we know which needs to be captured is like we call the density of the community or the amount of unity.

So, here we defined something which is called the density of a cluster. What is the density of a cluster? So, I will give you a one liner let us see. So, the density of a cluster is we say that the density of a cluster is D, if you look at every node in this cluster and at least D fraction of these nodes friends is in this cluster itself. It might be a bit difficult let us look at it with an example.

So, let us say that here is this cluster of 5 people. So, there are these 5 people in this cluster and let us say I say that the density of this cluster is 0.3 if I look at every person in this cluster and 30 percent of the friends of this person is in this same cluster. So, let us say that in this cluster everybody is having 10 friends.

So, I look at this node A out of these 10 friends 3 should be inside this cluster. For B out of these 10 friends 3 should be inside this same cluster and so on for the rest 3 of the nodes. So, that is what we mean by density of a cluster. So, I hope the definition is clear now.

(Refer Slide Time: 06:34)



Now, I am going to give you a very weird claim and the claim says that in this network where the threshold of adoption for every person is q as we said before. Then the cascade cannot complete itself, complete itself means starting from some nodes it cannot sweep the entire network, the cascade cannot complete itself. If there exist a cluster in this network of density greater than 1 minus q.

So, how trivial or non trivial is that? How do we prove it? So, the proof is actually very straightforward. We will let us look at it with our example, that will make it all the more easier.

(Refer Slide Time: 07:15)

For example what did it say that threshold for going outside is 40, percent threshold for staying in library and studying is 60 percent. And, then in this entire class having different, different clusters, these are those 2 people who decide that they are going to go outside and this idea of going outside starts diffusing on this network.

Now, assume that there is this cluster of density greater than 1 minus q. So, q is here 40 percent. Cluster of density greater than 1 minus q means that, there is a cluster of density greater than 60 percent. What does that mean, cluster of density greater than 60 percent? It means that if I look at a cluster here and look at every person here 60 percent of their friends are in the same cluster.

And we know that initially the complete class had actually decided to stay in library and work and here these two people have changed their mind and they said that we are going to go outside. Now this idea of going outside is diffusing through the network. And let us say while diffusing this idea comes close to this cluster. We are talking about and tries to enter this cluster.

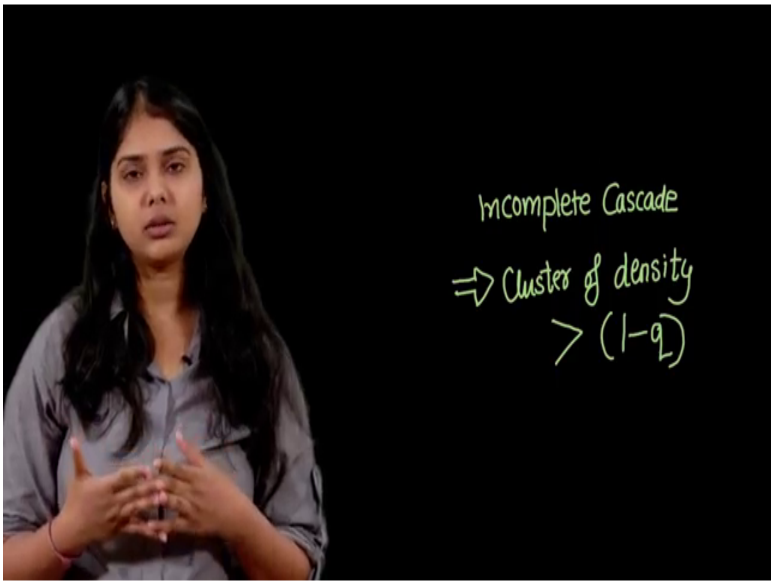
Now, see what happens. When this idea tries to enter this cluster, there is this node So, this node since it belongs to a cluster of density 1 minus q, density greater than 1 minus q that is density greater than 60 percent. 60 percent of his friends are in the same cluster; which means that 60 percent more than 60 percent of its friends have already decided to sit in library and study.

So, the threshold condition is not reached. For this node to go outside the number of its friends going outside, the fraction of its friends going outside and enjoying should be greater than 40 percent, but that cannot be achieved right. Because, 60 percent of its friends are already in the same cluster and have decided to stay in library and work.

Similarly, this happens for the second node of this cluster, third node of this cluster. So, this cascade is unable to enter inside this cluster and hence you see our claim is achieved. So, if there is this cluster having density greater than 1 minus q; no matter what this cascade which was occurring on this network cannot enter inside this cluster and hence cannot complete itself.

Now I am going to give you another weird statement and let us see how do we prove it. So, the first statement, the first claim which we made here was if a cascade is diffusing, if a an idea is diffusing on a network and if there is a cluster of density greater than 1 minus q, then the cascade cannot complete itself.

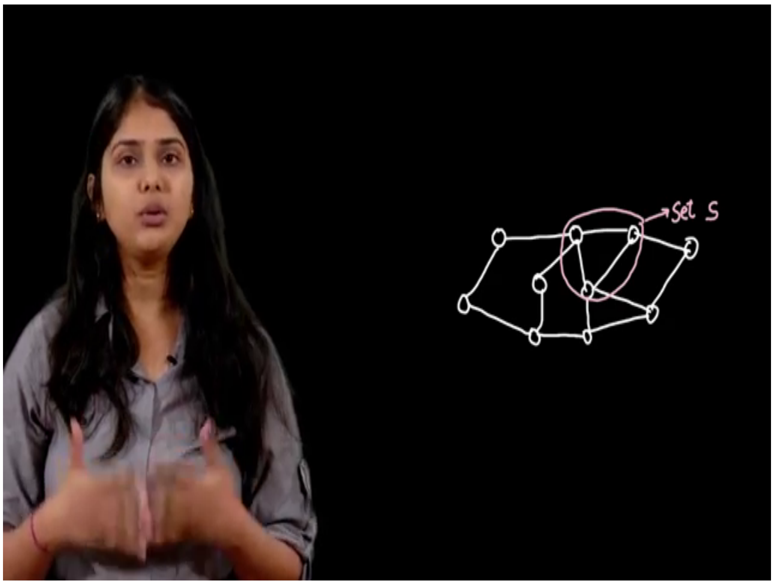
(Refer Slide Time: 10:27)



Now, what am I going to say? I am going to say that I give you a network and I give you the information about the cascade and I tell you that the cascade on this network is not complete. There is something which happened inside this network and the cascade could not complete itself, there are some nodes which did not adopt the idea; it means that if this happens that the cascade is not complete.

This implies that this shows that there is a cluster of density greater than 1 minus q in the network. So, it is a by implication, if you know propositional logic else we can just leave it aside. So, it is both ways. So, I leave it as an exercise problem to prove it. It is actually very intuitive like the similar way we did this proof. Let me give you a small hint.

(Refer Slide Time: 11:17)



So, when this cascade is diffusing on this network and we say that the cascade is not complete, it means that there are some nodes in this network who have not adopted this idea of going outside. Why did not they adopt this idea of going outside? Because, their threshold was not reached; so, let us represent these set of nodes as set S who did not adopt your idea.

So, these people who are staying in library and studying, we represent them as a set S and then one by one you examine this person. Why have this node not adopted the idea is because, its threshold is not reached, so using this you can prove this another side of the claim. So, I leave it for you to prove it.