Social Network Analysis
Prof. Tanmoy Chakraborty
Department of Computer Science and Engineering
Indraprastha Institute of Information Technology, Delhi

Chapter - 04
Lecture - 22
Lecture - 03

We have been discussing you know different ways to measure links ok and so far we have seen sign networks different ways to interpret sign networks and so on. Now let us look at other perspectives again let me repeat these are more of a social science kind of perspectives which we will try to quantify you know using network science theories ok.

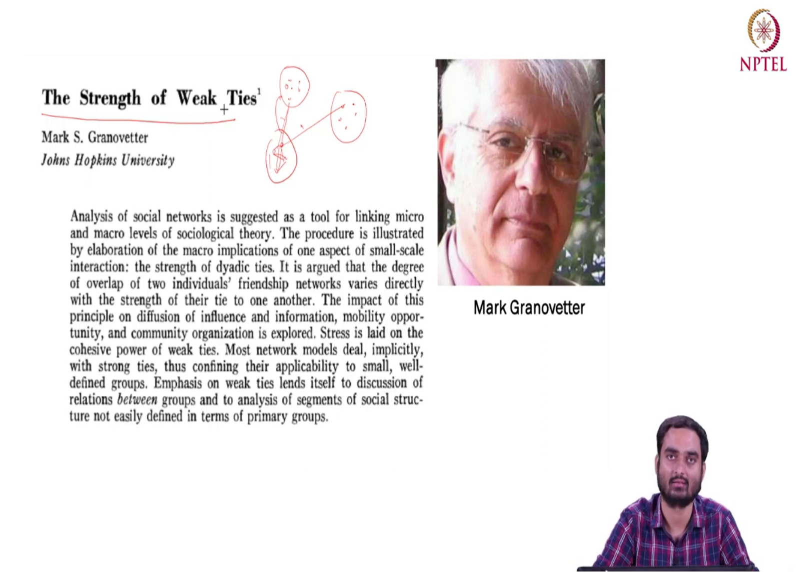
(Refer Slide Time: 00:47)

So, let us look at the interpersonal ties ok and we already mentioned about strong ties, weak ties and ties which are absent ok. So, strong ties or strong edges are those edges which are generally developed among entities that share interest and beliefs. For example, strong ties between right two friends, two individuals, who are part of same college for example. You can also thought of it as you know some sort of source of confidence or emotional dependency right say strong ties are often built among individuals within the same family ok.

Say between father and son right father and children right. Whereas, weak ties are generally developed between individuals who are not that connected, but you know they still share some commonalities right. For example say two there are two individuals who are who go to the same college right. But say after 10 years right they actually move to some other places.

Friend 1 will move to say Australia and friend 2 will say move to US and you know after 10 years they may not be able to talk to each other quite frequently. But since they shared some common interest long times back there is a tie between one and two right these are kind of weak ties. So, now why do we need to quantify strong ties and weak ties?

(Refer Slide Time: 02:36)



So, there was a very famous experiment again social science experiment way back 1970s, 1971 during that time by the famous you know famous anthropologist Mark Granovetter ok. So, Mark Granovetter what he did he actually interviewed you know few persons few individuals who recently moved I mean that time moved from one industry to another industry one company to another company. And then so basically Granovetter asked that why did you move ok and where did you get the information from right.

Did you get the information about the new company from your close friend, from your family friend or you got the information from somebody who you may not interact with quite frequently ok. So, he interviewed you know quite a few individuals who recently moved from one company to another company and it turned out that most of this movement happened right because of friends who were not in their strong in their strong circles right.

Say for example, you are currently with TCS, Tata Consultancy Services and you want to move right you want to move out you want to move to another industry. What would you expect would you expect that your other colleagues from TCS would give you information about new openings or you expect that colleagues who are outside TCS.

For example, your friends in Infosys, your friend in Wipro right. They would give you opportunities in their companies right or they would give you opportunities in some other companies. You can easily expect that friends who are not part of TCS would be useful in this case. Because they would give you know they would give you opportunities about you know places which are I mean which may be competitors of TCS for example, ok.

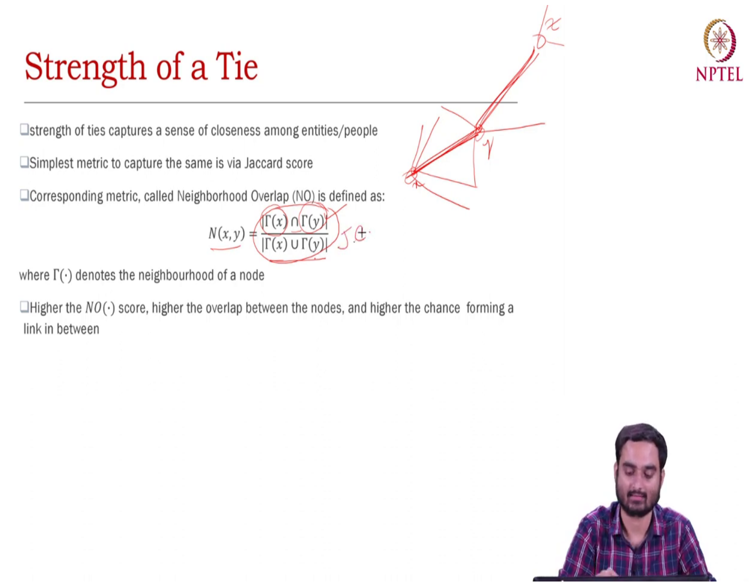
So, particularly from the recruitments point of view it is quite trivial to understand or it is quite trivial to you know believe that this kind of referrals essentially come from friends who you do not interact with frequently ok. And these friends are essentially your weak ties, your weak friends right.

Now if you think of; if you think of different clusters in a network right. So, say let us say; let us say these are yeah I mean each individual cluster. You can think of think of as close community and nodes interact within clusters very frequently. But there are links which connects two clusters and these links are essentially weak links ok.

So, if you want to really move from this cluster to this cluster right your weak links actually play an important role right. In terms of referral in terms of you know better opportunity better salary and so on and so forth. I mean it is not the case that your strong friends are jealous about you that is why they are not giving you a they are not I mean they are reluctant to give you information it is not the case. It happens because they do not know the I mean most of the times they do not they may not know the other opportunities right.

So, those friends which are not close by right those friends which are part of other companies they know the opportunities better than those who are part of the same community which you belong to right. Now this famous paper this is one of the seminal papers in social science network science the strength of weak ties. I think it has got more than 4000, 5000 citations I think till date. This paper in this paper it was explained again from social science point of view that why weak ties are important right.

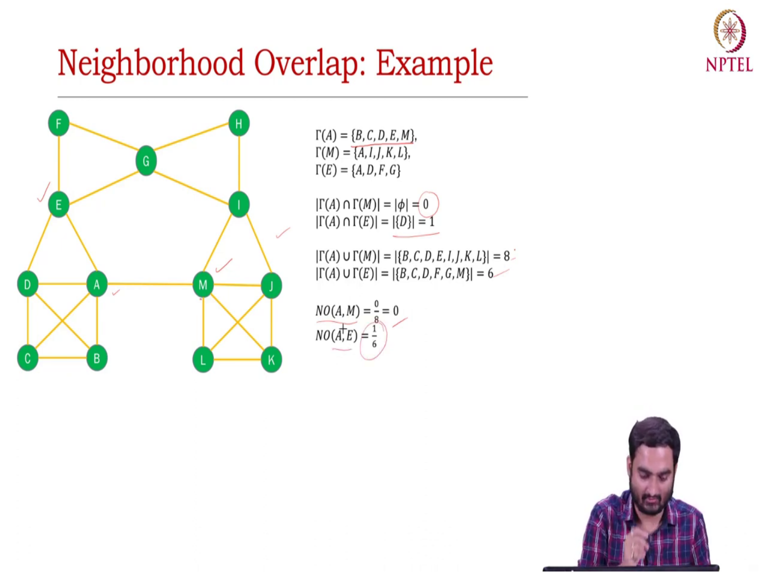
(Refer Slide Time: 06:44)



So, now how do we quantify the strength of a tie, strength of an edge? We I mean the simplest way to quantify the strength of an edge is you look at the neighborhood intersection of the neighborhood intersection of nodes I mean the neighborhood intersection of two end points of an edge right. Let us say x and y x is connected to some nodes, y is connected to some nodes right. And if let us say both x and y share many common friends therefore, it is highly likely that this edge x y edge is a strong tie.

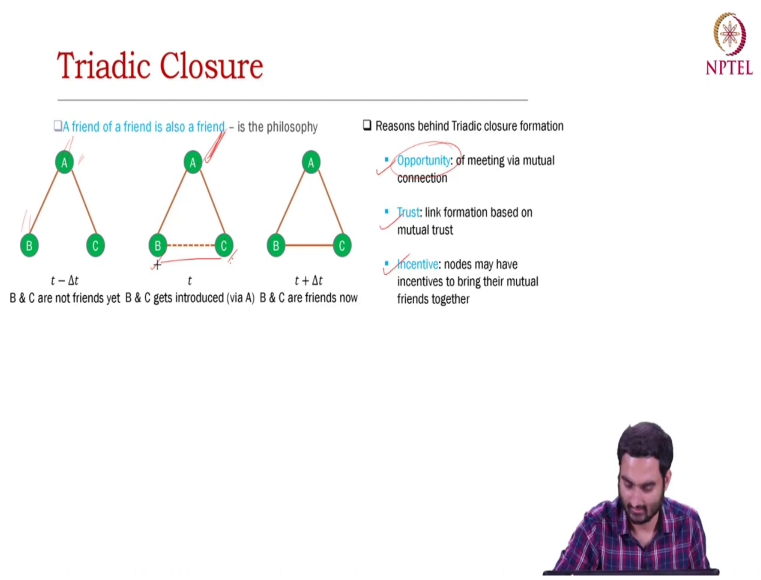
Whereas, there are there is an edge y and z, but none of this y and z you know share common friendships therefore, you know this might be a weak edge. So, the easy way to quantify it is to do some sort of intersection, so intersection of neighborhood and union of neighborhood. So, intersection is gamma x is intersection is a neighborhood set of x neighborhood set of y intersection and union this is called Jaccard coefficient ok. Jaccard coefficient of the neighborhoods of x and y alright.

(Refer Slide Time: 07:56)



So, if you look at this example here AM and E right. A has you know these friends B CDE BC DEM M also has friends and so on. So, the intersection of A and M is 0; there is no common friend intersection of A and E; A and E is 1. So and the union of these two things is basically 8 and 6 respectively. So, the neighborhood similarity of A and B; A and M is 0 and A and E is 1 by 6 ok. So, you can say that ok this is the weak tie right.

(Refer Slide Time: 08:50)



So, another similar concept is something called triadic closure we have already discussed before triadic closure measures that ok. If x and y say in this case if A and B are connected and B and C are also connected. A and B are connected and A and C are also connected. What is the likelihood that B and C will also be connected in the future ok? So, the more the triadic closures you have in a graph the higher the clustering coefficient.

The clustering coefficient is a way of measuring the triadic closure. Now let us try to understand different reasons behind the formation of this triadic closure. So, one reason is the opportunity right. So, every mutual friend in this particular case say mutual friend is A right. So, every mutual friend like A between B and C you know gives. So, A will give opportunities to both B and C to meet and become acquainted right with each other thereby increasing the chance of developing a connection more of an opportunity ok. So, let us say there is a referral system ok.

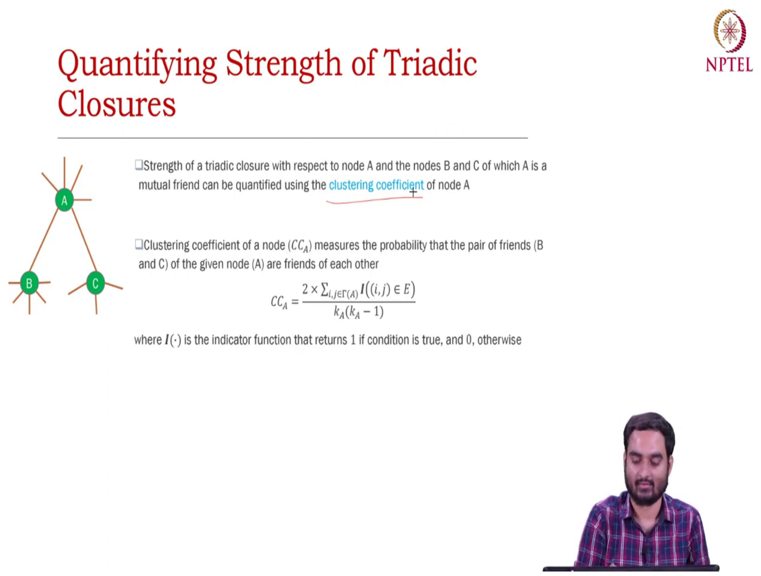
So, A knows that B and C are good friends, but if you refer to B and C to become friend A will get some incentive ok. I will discuss incentive in the later part ok. The next one is more of a trust we discussed sign network. So, in a network if connections are formed based on some sort of mutual trust ok then x right in this case B and C right; B is more likely to trust C B is more likely to trust C. Because C has already trusted A and B also has already trusted A.

So, due to this balance theory social you know status theory kind of theories a trust over the relationship can be formed between B and C. So, the other the third reason is incentive ok. So, let us say A has an incentive to bring its mutual friends B and C together in order to increase the interpersonal trust and cooperation among peers right.

Therefore, I mean this trust and cooperation would help A right. Let us say you know A is a part of a company again B and C also part of a company and A is kind of a manager and if A feels that I mean both B and C are reporting to A right, but B and C are not talking to each other that frequently.

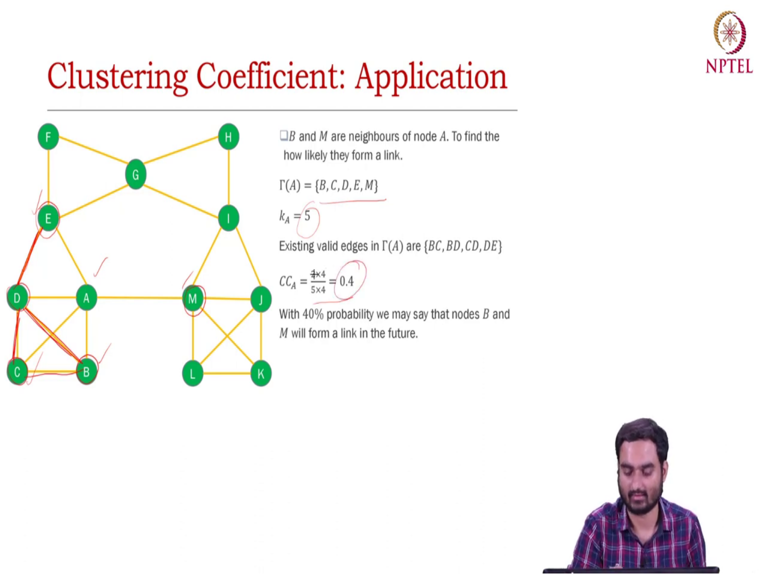
So, what A would do? A would try to convince B and C that hey why do not you talk, why do I mean why do not you talk to each other and you know start collaborating. And through the collaboration possibly the trust and you know the cooperation and trust will increase overall in the organization and we will get some benefit ok.

(Refer Slide Time: 12:09)



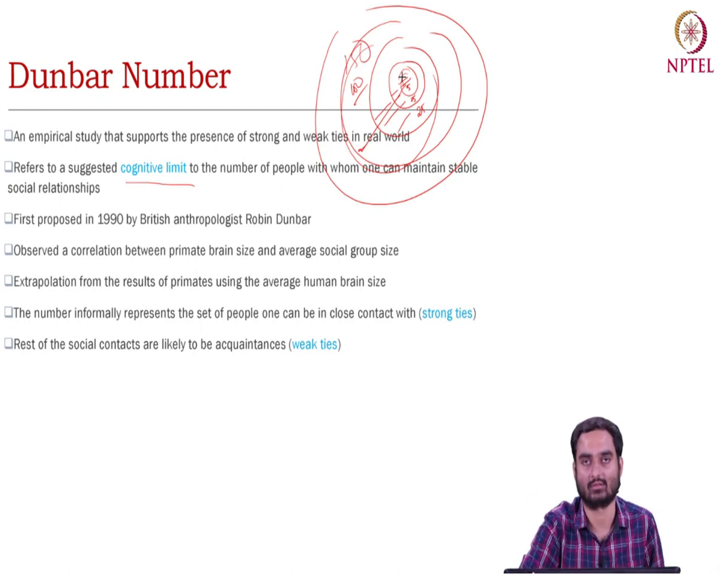
So, these are the reasons and how to quantify a triadic closure. We have already discussed. We can use clustering coefficient, local clustering coefficient which is basically the actual number of edges among the neighbors divide by the possible number of edges among neighbors right.

(Refer Slide Time: 12:30)



So, if you look at here in this example say in case of node A, node A has degree 1, 2, 3, 4, 5 and the clustering coefficient and among them this is sorry this node, this node, this node, this node and this node. So, this edge exists this one also exists, this one exists, this one exists right that is all there are 4 edges which exist so and there are 5 nodes. So, 5 C 2 pairs and 4 pairs are connected therefore, the clustering coefficient is 0.4 ok.

(Refer Slide Time: 13:13)



So, now let us look at another interesting social study. Now this is called Dunbar number ok. What is Dunbar number? So, Dunbar in 1992 he tried to understand the relationship between the cognitive limit, cognitive limit of an individual to remember the people or friends who the person interacts quite steadily quite frequently ok.

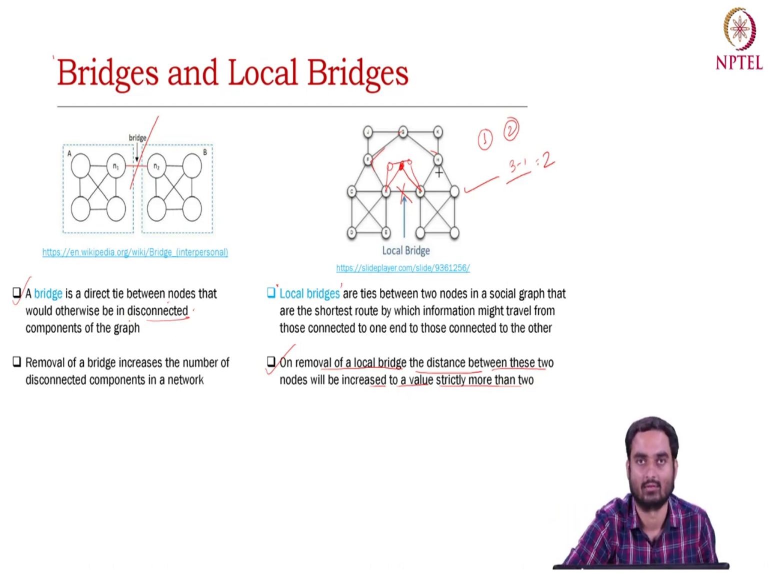
So, let us say; let us say you have 2000 friends on Facebook right. But you may not be able to remember each and every friend each and every friends activities right. You may not be able to remember by name in fact, all 2000 friends right. And what Dunbar mentioned Robin Dunbar mentioned that you know our brain is capable of remembering only 150 close friends ok.

So, if you think of; if you think of you know say you know encircling friends based on the cognitive based on your cognitive ability. You will see that there are you know you can think of this kind of circles right this kind of circles where say there are only 5 friends who you interact very very frequently.

There are around 15 friends who you interact little bit less frequently there are 25 friends who you interact even less lesser frequently and so on. And you see that there are you know in total there are 150 such friends which you can remember by name or by their activities, the remaining friends you may not be able to remember ok.

And this Dunbar’s number has a direct resemblance with the with the code periphery structure that we discussed in the earlier chapters right. If you think of if you focus on a particular node and look at the first core right first of neighbor, second of neighbors, third of neighbors and so on and so forth. You see that after certain points you know the central node right which is the ego node. The ego node will not be able to remember more than 150 alters ok.

(Refer Slide Time: 15:47)



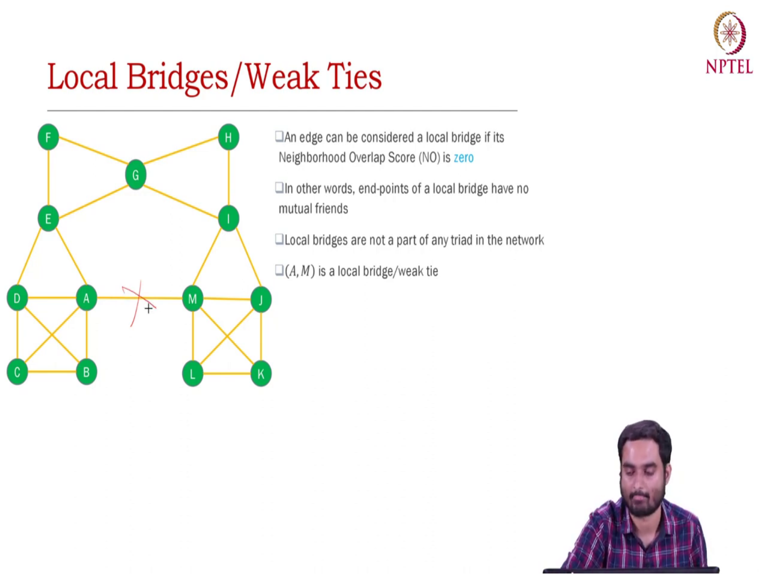
So, this was a very interesting study and you know you can also try to quantify edges based on something called bridge, local bridge and so on. So, basically what is a bridge? A bridge is a tie or an edge between nodes that would otherwise be in disconnected components of the graph. Now if we remove edge if we remove the bridge you would get you would basically disconnect a network ok. So, this is also related to weak ties, but you know bridge is a very strong definition right. So, can we reduce the constant? Now so you basically define something called local bridge.

So, a local bridge is basically a tie between two nodes whose deletion would increase the shortest path distance between two end points. Let us say in this case right say A and B now A and B are directly connected. If you remove this 1, what would happen? The shortest path between A and B would be 1, 2, 3 and 4. So, there is an increase of shortage path. Now what local bridge says? So, on removal of a local bridge the distance the shortest path distance between two end points will be increased to a value strictly more than 2 ok.

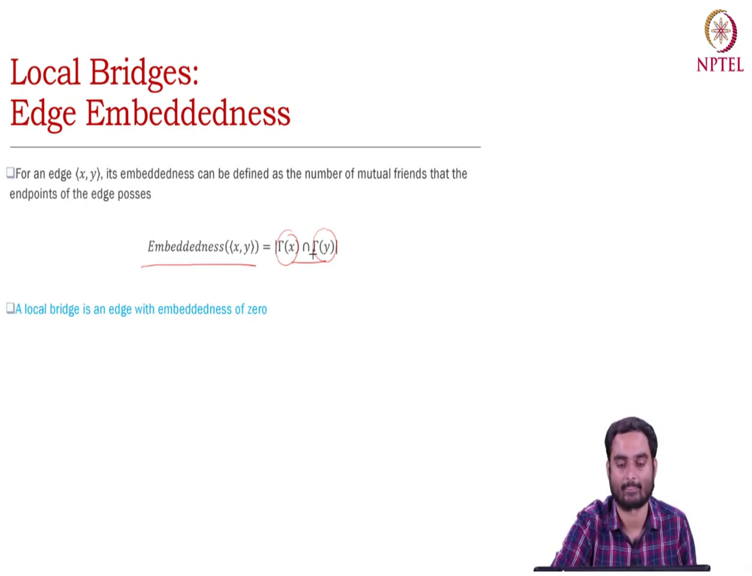
So, let us say there is a connection like this there is a node here right. Now if you remove this what would happen? The distance between A and B is still 2 earlier it was 1, now it is 2. So, the increase is not more than 2 therefore, this is not a local bridge right. Let say there is a connection like this if you remove this, what is the distance is 3. So, 3 minus 1 it is 2, so still it is 2 right not strictly more than 2, but 2 is ok.

So, what you are trying to understand? You are basically trying to understand whether there exist there exist other mutual neighbors between A and B say in case of this kind of structure right. So, this is a mutual neighbor of A and B and if you have a mutual neighbor right then there is no point in deleting it because through the mutual neighbor they will be connected ok.

(Refer Slide Time: 18:31)

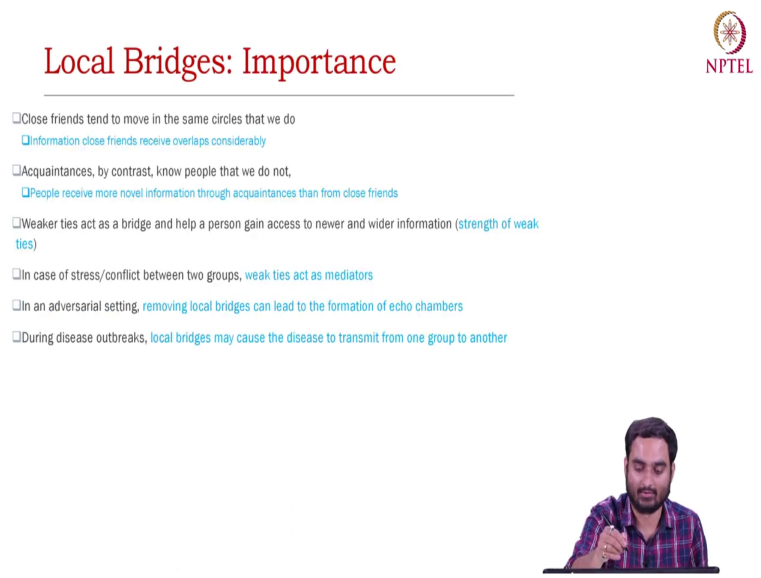


(Refer Slide Time: 18:43)



So, here you see that AM edge this is a local bridge alright. So, how do we measure local bridge? There is something called edge embeddedness embeddedness is very simple basically neighborhood intersection. You look at x's neighbors and y's neighbors and if the intersection is zero. Then it can be qualified as a local bridge otherwise not ok.

(Refer Slide Time: 19:03)



So, what is the importance of a local bridge? Importance of a local bridge is that it basically indicates whether these two nodes are very close by they receive new information novel information right. Whether removing the bridge will disconnected it right and whether this bridge can act as a link which helps you which helps the information move from one community to another community ok.

So, this was about local you know bridge Dunbar's number and you know other aspects of link analysis. So, we stop here and in the next part we will discuss page rank and other quantification of link.

Thank you.