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

Chapter - 07
Lecture - 01

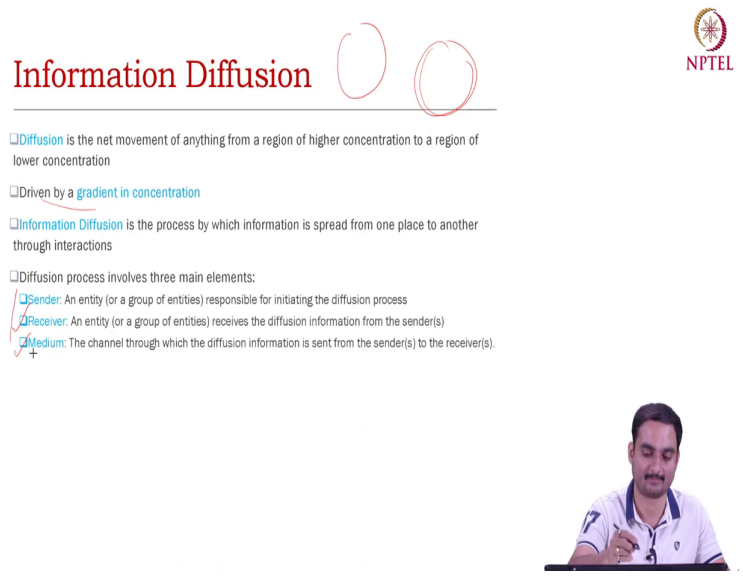
Welcome to the 7th chapter of the Social Network Analysis course. And today we will discuss a very important as well as relevant chapter, relevant I mean this is pertaining to our I mean you know current time, where we are actually at the middle of Covid-19. So, we will discuss how information spreads over social network and how disease the epidemics different kind of epidemics spreads over you know user network.

So, this chapter is on cascade behavior and network effects, essentially we will study you know on social network how information like propaganda, misinformation, fake news, agenda, right diffuse over time. We will discuss how users adopt a particular say product or particular agenda, particular behavior over time and how you know; I mean what is the effect of social network in general.

What is the effect of neighborhood structure of a user which basically motivates the user to adopt a particular say product or opinion, ok. So, in this chapter what are the major learning objectives we will discuss you know a wide variety of models for information diffusion. We will discuss how this diffusion models you know evolved over time, you know we will look at models which are proposed long time ago.

In fact, those models were not related to social network, when those models were proposed. For example, epidemic models right those models were proposed, those are mostly those mostly come under the broad area of you know statistical physics, but one can also adopt those models for information diffusion. We will also understand the you know the advantages and limitations of different models and we will try to think how we can; you know how we can design our own diffusion model right, which can overcome some of the limitations.

(Refer Slide Time: 02:44)



So, diffusion is basically a kind of a movement of anything right, it can be a movement of an opinion, it can be a movement of a propaganda for example, or a movement of fake news right, movement of anything right. And this movement basically happens due to the interactions ok. For example, on social network, when we re-tweet some other tweets, when we share some other posts we basically adopt to that particular thing and then we further move that entity right, to the next step right.

So, when I say when I write a tweet, all my followers will actually exposed to the tweet right. Now, out of say n out of n number of followers say n by 2 followers would further retweet my tweet. So, it means that my tweet now has moved from one node to the other node ok, or one layer to the next layer. Now, next say for example, one of their followers will further retweet it. So, in this way the information moves from one node to the next node to the next node, ok.

And in case of offline social network, say for example, user interaction network, say if you think of some sort of contagion right, say again Covid-19. Now, it basically spreads due to the contact, due to the interactions close interactions right when two individuals come closer and say one individual is already affected by Covid-19. So, with certain probability the other user, the other individual would also be infected by Covid-19 and so on and so forth, right.

So, in the epidemiology, in the disease spreading behavior, if we think of you know what would be the expected number of users at certain point in time, who are going to be infected

by this disease right; how can we how can we design such predictive models? I am pretty sure in you know last 1 or 2, 1 or 2 years we have seen many such models, many such information diffusion or Covid-19 diffusion, Covid-19 spread models to predict you know number of deaths, number of infections and so on and so forth.

So, you may wonder how you start right, how do you start from a particular model what would be our base model and how do we improve this in further ok. So, we will discuss all these nitty gritties in this particular chapter. So, as I mentioned basically depends on the gradient in concentration, now say for example, in a particular sub graph of a network right, you see that you see that most of the users are already infected right, within a community most of the users are infected.

Now, what would be your optimal strategy? For example, when it comes to you know vaccinate individuals right; what would be your optimal strategy so that the information or so that the in this case infection would not be able to move from say this part of the network to the this part of the network right. And say for example, you do not have the, you do not have the liberty, you do not have the enough, you do not have enough fund to vaccinate each and every individual.

So, what would be your optimal strategy to choose nodes, right so that choose nodes and vaccinate them, so that the you know the infection will not move further, ok. So, information diffusion is basically a process by which information is spread from one place to the another through interactions ok.

And these three terminologies we are going to use in this particular chapter, there is something there is some entity called sender who would basically would be responsible for sending that information or the disease. Receiver who would receive it right and then the medium, medium is basically the channel through which this particular information is sent from a sender to a receiver, ok.

(Refer Slide Time: 07:17)

It has a lot of real applications, as you may you know easily understand. Health care, disease propagation, epidemic spreading, modeling and so on, right, there are infact models to you know to mimic the way Ebola spread long time back. Similarly, we see Covid-19 related models and so on. In socio political cascades it is very important, how do we how do we model that say whenever a government right, announces some say some decision, right.

How this decision will be perceived by individuals, whether how people will react to that particular decision, whether they would support they would oppose, they would further move this thing move the decision from one node to another node and so on. In fact, there are also studies, which basically showed that during this political movement, how recruitment happens.

Say for example, you want to you want to, you know run a particular political movement right on social networks and you need to recruit individuals, so that they further, you know they further move your propaganda to other individuals, right. So, the question is how do you recruit? Now, this recruitment the I mean the act of recruitment this also comes under say other kind of recruitment process.

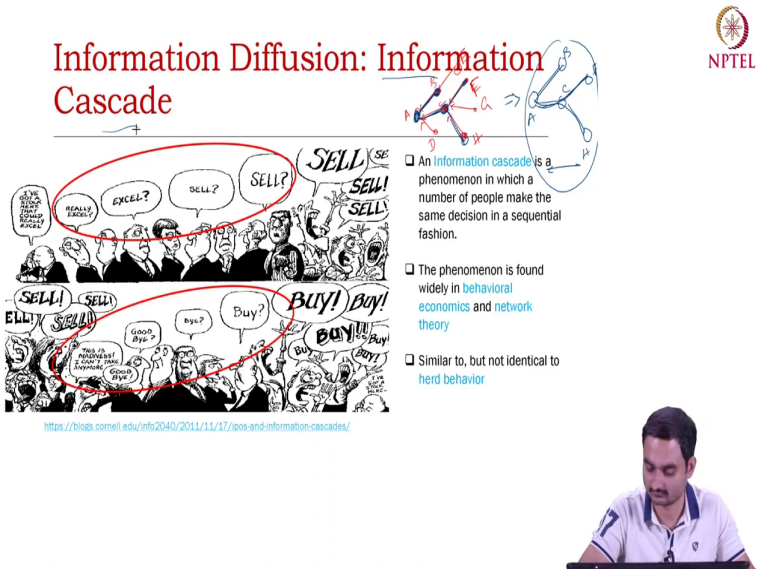
For example, terrorist recruitment, right. We have we may have heard that in online social, through online social network you know criminals, terrorist are also being recruited and they also follow certain strategies. It is not like a random recruitment right, they follow certain strategy and through the strategy they recruit individuals over time.

We of course, must have heard of this movement called MeToo right, it is MeToo MeToo movement was you know that got kind of viral right and not only in India, but also outside India. And following this MeToo movement there are other like, StarToo movement, right HimToo and other types of movements also happened. And people started talking about their you know their sexual abuse and other harassment kind of activities and how they got victimized, right.

So, in financial market cascades also we talk about stock market right, how certain stock becomes so viral overly popular among investors, how viral marketing happens. In fact, when you talk about targeted advertisement right, say for example, you want to you want to say you want to advertise your product right, but you do not have that bandwidth, you do not have that money or fund to you know to look at all possible individuals in a social network, right.

So, how do you choose again individuals or groups of individuals in a targeted manner so that your marketing strategy is maximized, right. Of course, in social network rumors spread, belief spread, propaganda spread, fake news, hate speech spread, offensive content spread, in general these are often modeled or understood using cascade behavior ok.

(Refer Slide Time: 11:01)



So, information cascade: now when you talk about information diffusion, information cascade and information diffusion these are kind of synonymous, they are used interchangeably. But when we talk about information cascade, we particularly look at how the tree like structure grows over time. What do you mean by tree like structure? Let us say,

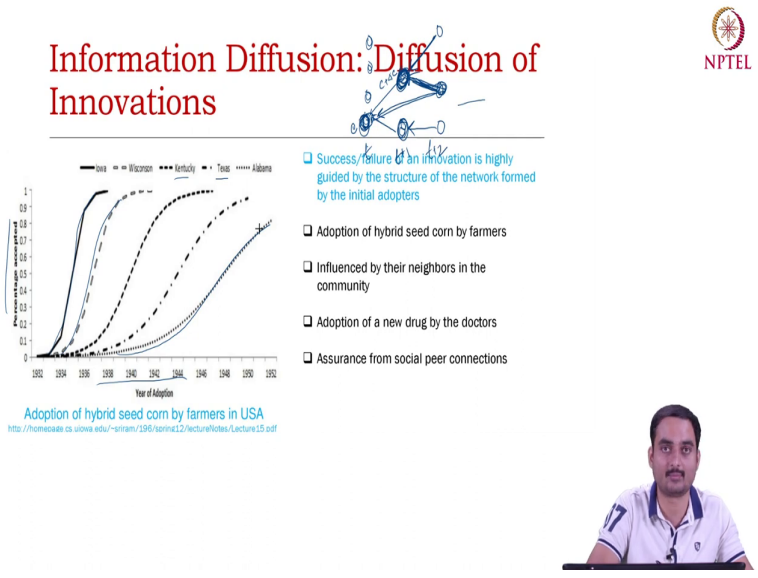
let us say right, let us say this is A and A is followed by B, C, D; B is followed by E or F and G, H, right.

And say A has retweeted something, right. So, among B, C, D who say let us assume that B and C have retweeted S to it, right. Then say from C F and H further retweeted right. So, if you look at only the nodes and edges, which have got affected, right. So, you will see that it basically follows a tree like structure. So, this is the source node then, right ok.

Now, from this twitter network you can create a tree like structure like this, right. Now, this is called information cascade. So, generally information cascade is modeled through a tree, right, but you can also look at, I mean if you do not want to model it as a tree as a tree, you just look at the individual branches right. For example, this branch is of size length 1, this branch is of size is of length 2, this branch is of length 2 and so on and so forth.

So, your target would be to you know to predict the structure of this tree, meaning that you can basically predict the depth of a tree, the breadth of a tree and you know individual branches and so on and so forth. So, if you can predict the tree right, then you are done ok. So, this is actually about information cascade right.

(Refer Slide Time: 13:44)



So, in fact, there is another interesting application of information diffusion which is the diffusion of innovation right. For example, let us think of a citation network right, and let us say let us say at time t this 4 4 papers, have been published and this paper has proposed an

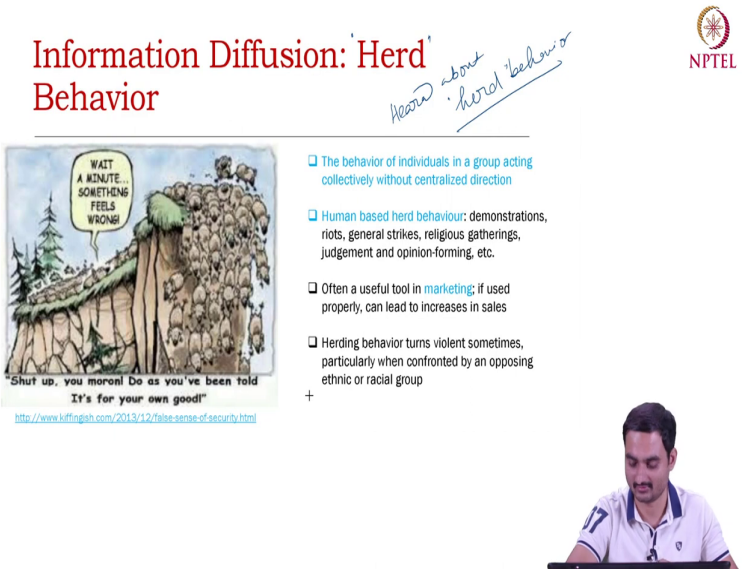
idea and concept and a new concept C right, and say these 2 papers at time t plus 1 have cited this paper right.

So, therefore, the concept has now been adopted by these two papers. So, the innovation has moved from this node to the to these two nodes right. Now, at t plus 2, let us say this paper has cited this paper right, this also cited this also cited this ok, right. So, it means that the innovation actually moves from this node to this node, also this node to this node.

It may happen that at this node there is some additional innovation which was moved we which basically moved from this node to this node through this edge, but this innovation moved from this node to this node directly, ok. So, you can also come up with a DAG kind of structure, Directed Acyclic Graph why it is a cyclic you think about it citation network is always a cyclic ok.

And it is detected because of the citation; I mean the directionality of the citation ok. So, in fact, if you a if you can think of the number of adopters right, versus year you see you can see these kind of patterns, you see here there are some sort of innovations right and thus those innovations were adopted by different cities. And you see for different cities like Iowa, Texas how this number of adopters changed over time ok.

(Refer Slide Time: 16:03)



So, you also must have heard something called Herd behavior right. So, this herd behavior herd immunity we have all heard in this Covid-19 time right, we have all heard about herd

behavior ok. So, what is this herd behavior? In general, I mean in terms of herd immunity, it basically says that you know the an individual is already immunized, right.

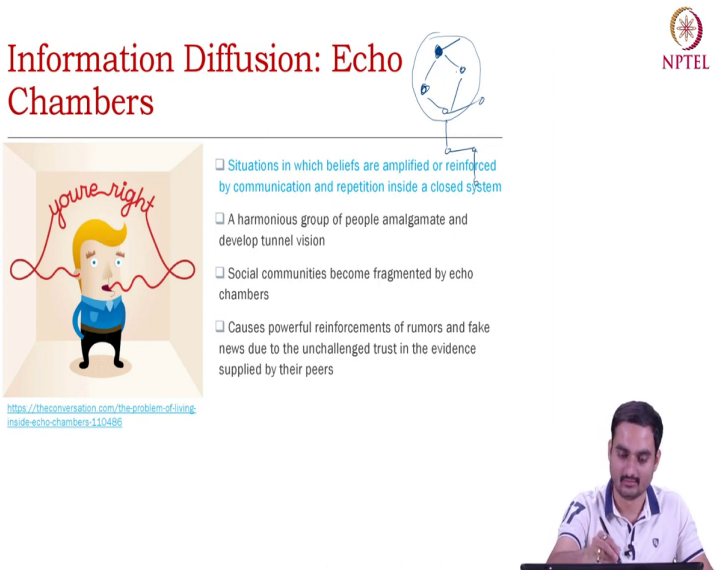
So, if he or she will be exposed to the particular virus he or she will not be infected, because of this herd behavior. Or you know over time the individual has got infected so many times that it is less likely that in the future when he or she will expose to the will be exposed to the particular virus, he will not be infected further ok. So, this is called herd behavior.

Now, in case of you know cascade behavior etcetera, when we say that when I look at my individuals right, and I see that most of my individuals have already adopted a particular behavior, it is highly likely that I would also adopt it, ok. So, this is kind of a herd behavior, ok. In fact, what happens is that due to the herd behavior, even if the rest of the individuals. In fact, this herd behavior also is related to my mental state right.

For example say there is a there is a chance that I can subscribe either, I can subscribe to say Amazon Prime, I can also subscribe to Netflix right, but since I belong to the US, right it is highly likely that I will subscribe to Netflix, not Amazon Prime right. Because in US majority of the individuals subscribe to the Netflix not Amazon Prime therefore, I am although I am an and say let us say in India it is highly likely that people generally subscribe to Amazon Prime.

So, even if I am an Indian, when I move to US right it is highly likely that I will subscribe to Netflix. So, this is kind of a herd behavior ok alright.

(Refer Slide Time: 18:34)



So, another important concept is called echo chamber. What is echo chamber? Let us say there is a network right and say you are a you are approved you are a democrat supporter, a pro democrat right. So, this is you and you are a democrat supporter and most of your followers generally retweet what you tweet, they are actually your genuine followers ok.

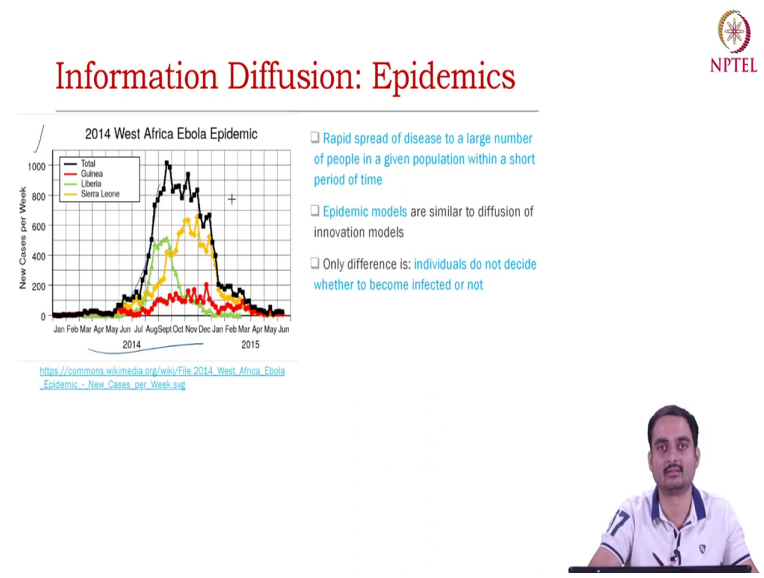
Now, what happens is that over time, when you tweet about tweet about some democratic policy or democratic what is you know next stage or whatever your, followers should also retweet it right. And since you are part of a community your other nodes in the in that community would continuously get exposed to the tweets that you tweeted or your followers retweeted, right.

So, what times what happens is that the size of the community right grows and the and I mean you can think of this community as not as kind of a structural community that we are generally talk, about topology based community. Now, these communities basically such a community which depends on how people react, how people you know retweet and so on and so forth.

So, what times what happens is that due to the echo chamber behavior right, you see since you have you are part of a community which are which is majorly pro-democrat community, what happens is that over times you would actually get exposed to more and more pro-democrat tweets right. And you tend to retweet those kind of tweets again and again ok. So, in other ways you see a kind of a kind of a filter bubble, it is also called filter bubble in some other terminologies, right. It basically starts increasing ok.

And it is called echo chamber because a particular idea gets echoed by individuals within the community ok. So, in fact, eco chamber is a very important concept in social science theory, where we see that how you know how people how all this fake news misinformation kind of stuff gets viral, gets get viral over time and how eco chambers are responsible for the spread of such offensive content ok.

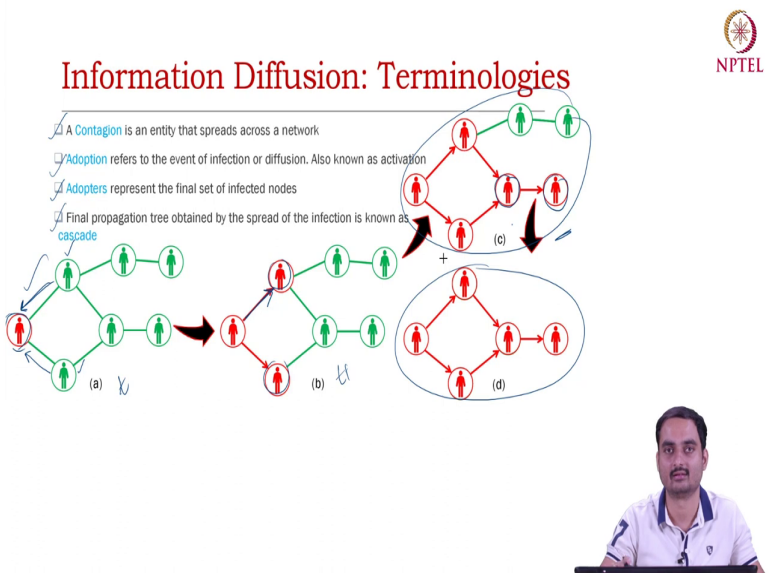
(Refer Slide Time: 21:14)



So, and of course, in the context of epidemics right, we now this is a study about how Ebola right spread in West Africa right, and you see that over time in different parts of West Africa you see the spreading the spreading pattern right increases. Of course, you see most of the cases one peak, then it decreases over time some sort of normal behavior, but it may not be the case every time.

For example, in Covid-19 case we have already seen 2-3 peaks right, in last 2-3 years. So, it depends on the pattern of the virus right, it also depends on government strategy for example, lockdown say arrangement of vaccines and so on and so forth. So, we will discuss about this kind of models later ok.

(Refer Slide Time: 22:14)



So, let us look at some of the some of the terminologies, which we are going to use in the remaining part of this chapter. So, what is contagion? Now, contagion is an entity that spreads across the network, this contagion can be a disease, virus a contagion can be an information ok.

Adopter: what is adoption? Adoption is basically referring to the event of infection or diffusion right, it is also called activation so, if I adopt something meaning, that in case of epidemiology I have already got infected, right. In case of say product purchase, basically I adopted a particular product right. In case of viral marketing I actually activated the particular product or campaign that is spreading over time, over social network.

Adopters are individuals who have adopted right. And of course, the final thing is propagation, basically a tree through which an information spreads over time. You see here say, this is the this is the sender right or initiator and this is at time t at time t plus 1, both the followers right have adopted that particular ideology. You see here the direction is in the opposite direction of the follower following. Now, since these two individuals follow this guy. So, this is the follower link right, but this link indicates the spread.

So, the spread happens in the opposite direction of follower, followed relationship, right. So, and now you see, you see this got this individual has got accept affected, then this one and so on. Now at so and then you see that you know things got stabilized so, from this network you can create this kind of cascade, cascade tree or whatever DAG right, kind of structure. So, we

stop here, in the next lecture we will talk about different types of information diffusion models.

Thank you.