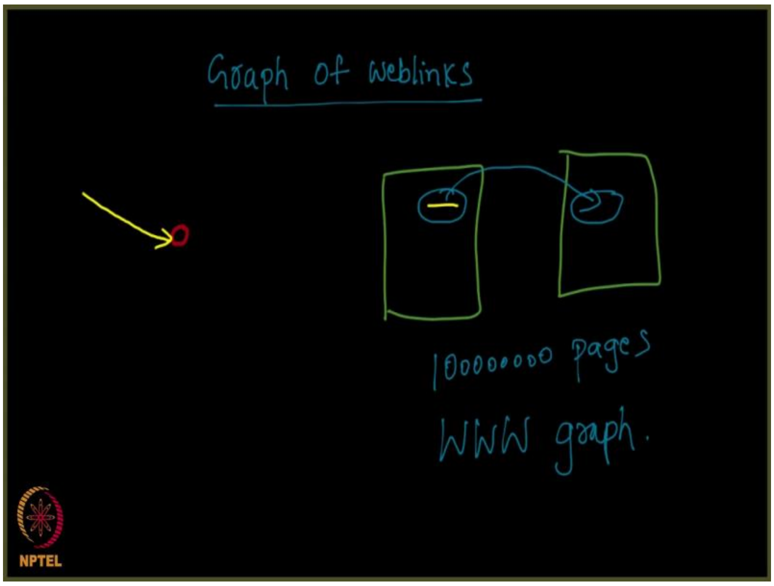
Social Networks
Prof. S. R. S. Iyengar
Department of Computer Science
Indian Institute of Technology, Ropar
Rich Get Richer Phenomenon

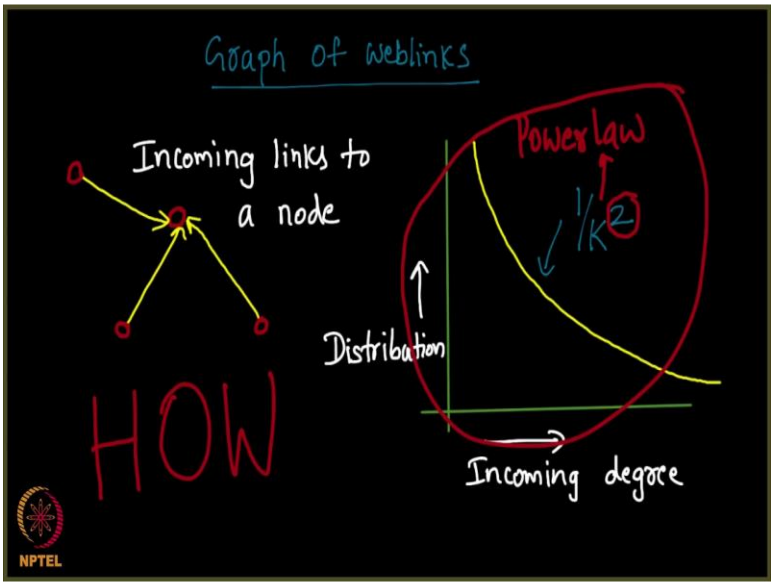
Lecture – 117
Power Law emerges in WWW graphs

(Refer Slide Time: 00:05)



So, remember that the graphs that we discussed before. If you do not remember, let me revise it quickly for you all. A web graph is basically a graph with a node ok. This node represents let us say a page like this a page pointing to another page right. So, the here is a page which has a hyperlink to another page. You click on this, you are taken to a new page right and I and internet has millions and millions of such pages right; several such a page, a whole lot of such pages correct. And then I look at the underlying graph and that is called the web graph, the graph of the World Wide Web.

(Refer Slide Time: 01:03)

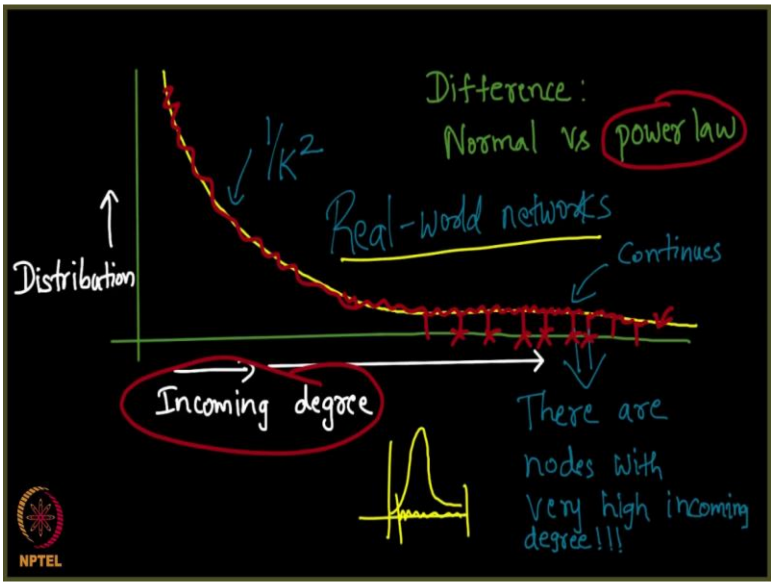


So, I am going to consider that right now ok. So, typically a node represents a node represents a page, this represents a page. And an incoming link to it represents a link that is pointing to this page as simple as that right ok. So, there may be many such incoming links to a webpage from some other page. By that time in if I put a link to your home page from my home page then, I put an edge from my home page to your home page. Home pages are all nodes here and if there is a link from one page to the other, you put a directed edge right. These are all the web pages from where there is a link put to this page.

So, these are the incoming links to a given right. Here the incoming link is incoming degree is 3 right; that is called the degree. Now, let me plot on x and y axis the following. On x axis, I will write the incoming degree and on y axis I will draw the distribution right and I will observe what is happening here. I observe that there is a drop like this. Unlike in the previous case, we do not see a normal distribution here and in fact, a closer observation even tells me that this plot very much resembles $1/K^2$ where K is the x axis. It resembles this right. And this is called the power law right; the law where the, the law which states that some distributions do not exhibit normal behaviour. You do not see a bell curve there; you see a drop like this ok.

So, the exponent here is to, it can be anything for what we now it can be anything. But in this case in the case of the web links, it turns out to be K square ok.

(Refer Slide Time: 03:09)



So, what are we observed so far? When you plotted the incoming degree versus the distribution right, you got this for web graphs. But then when you would you this is 1 over K square, but then there is a difference. What is the difference? Normal versus power law.

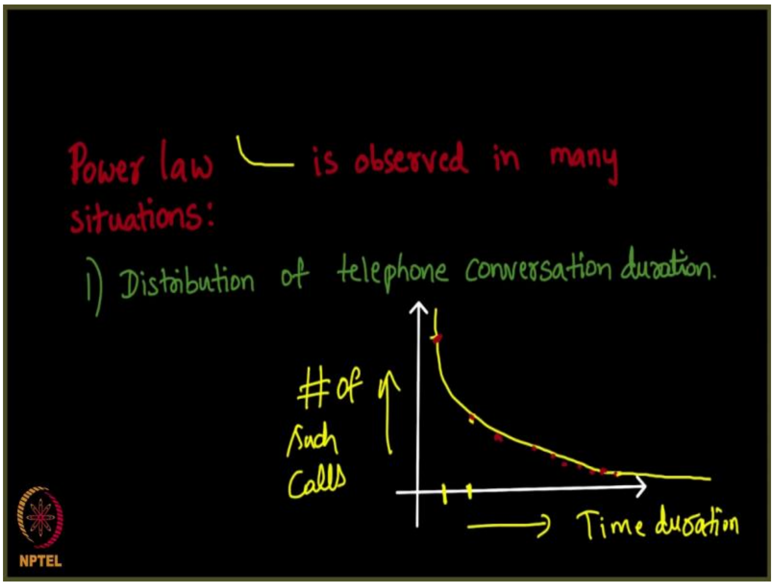
What is the difference? See, initially you saw we plotted two graphs. One was normal, one was power law right. What is the huge difference between these two things? What does power law tell us which normal distribution did not? We see this in fact, this actually extends right. The this is the yellow line here, let me show that you once again; this actually extends, it goes beyond; for a for a long time you will see a curve like this ok.

So, let me just extend this green line also. And what is this signify? What is this band here? What is this band here? What is it denote? Please look at the x and y axis, its rather self explanatory. What does it denote? It denotes something very significant, observe carefully. It denotes that this continues this goes on; goes on means what? It means, there are nodes with very high incoming degree Do not you think that is what this means? Look at this side this x axis, x axis means what?

We just note it, the incoming degree. The incoming degree is very high here as it goes on the right side of the x axis and the yellow line denotes that there are nodes; there are nodes with very high degree. That was not the case in normal distribution as you remember correct. The power laws simply say that the curve goes a extends beyond a stage. And it sort of is close to the x axis although it does not become equal to x axis; it goes on which means the it shows of shows us of the existence of nodes that have very high incoming degree right, fine. Here is the question. Why do you observe this drop of the curve? And secondly, how come there are so many nodes with higher and higher degree is observed in real world networks?

I hope you remember the in the previous lecture. One of the previous two lectures, we observed that if you plot it for let us say random graph which is a GNP model, with some 1000 nodes and probability 0.1. You observe that it was a quick drop. How was the graph? Remember the graph was something like this that is all. After this there was no extension, before this there is no value no node which satisfies this. It is all it all lies between this and this, but in case of real-world networks we see an its sort of extends. Firstly, it drops it drops and extends right. Both funny and strange; why is this happening will be the crux of this chapter; and the crux of discussions in the forthcoming lectures. Let us go ahead ok.

(Refer Slide Time: 06:39).



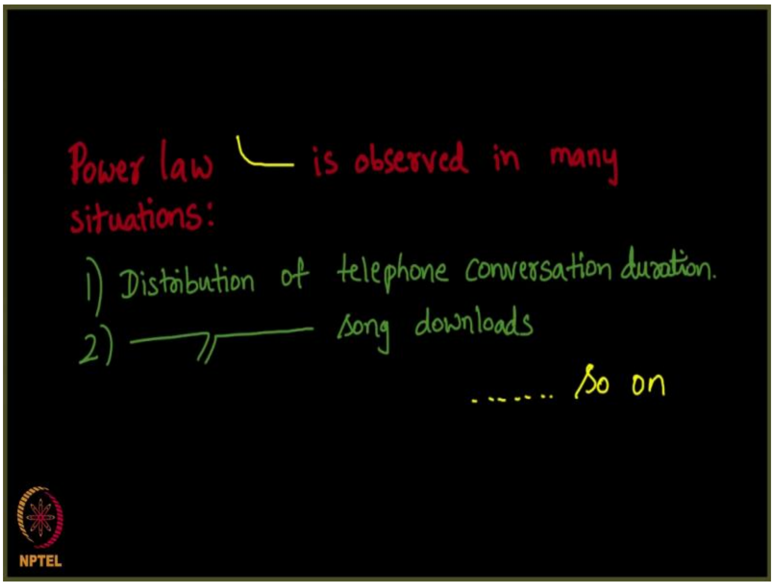
So, it has a very high degrees what we observe remember that. There are many nodes with very high degree it goes on and on. This is called the power law. Power law is basically the drop of the plot the distributional drops ok. And that, that resembles a $1/k^2$ or k^3 or in general $1/k^{\alpha}$. Let us say all right.

So, the power law is a drop, it is observed in many situations, not just in the in. What did we observed in the previous case? We saw the distribution of let us say the incoming nodes of web graph. And we observe it is a power law then, immediate next question you may want to ask is where exactly I see power law apart from these web graphs. In fact, you see them in many situations. You will see them in telephonic conversations let me go back so, that I can show them to you one by one. So, you will see this distribution, this kind of distribution is observed in the telephonic conversation.

For example, what does this means? This means you again plot x and y axis. On x axis you plot the time duration of phone calls and on y axis you plot the number of such phone calls or percentage of such phone calls. Here both are one and the same. And you will observe that it is again I drop like this ok. So, what do I mean by this? How many phone calls have been of 5 minutes duration of whole lot of them? How many of them have been 10-minute duration of whole lot? But relatively less compared to what used to happen for a 5-minute phone call duration. That is a whole lot as you can see 10 minutes is less 15 minutes is a lot less so on and so forth right.

And you do find this extension; you do find there are phone calls which are of a high time duration right. This distribution is also a power law very surprisingly ok. Where else do we see power law? We see that in many places. In fact, this data sets are available online, you can just check.

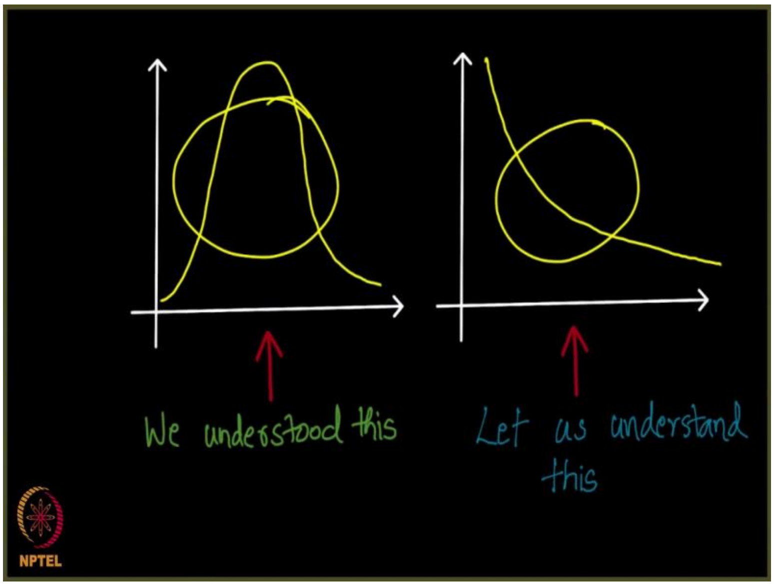
(Refer Slide Time: 09:04)



So, where again do we see the distribution of song downloads right? What are the numbers of songs? What are the songs which are downloaded more than 10000 times from our website? There are some official a song downloads website right.

So, if you go and look at this even these shows a power law ok. This was a exponential drop lines ok. So, which means power law is observed in many places. While we think, it should have been a bell curve in place of bell curve. There is a drop like this ok. So, there are many such networks you can take a look at it. In fact, there are hundreds known till date or may be even more right. I i know at least couple of dozens of a very well-known network that exhibit power law. And people are even looked at why exactly they are exhibit power law which will be the focus of our next lecture.

(Refer Slide Time: 10:02)



So, now we observe that there is a normal distribution. There is a power law distribution right. So, this is this we completely understood right. We understood it completely why was it? Remember I took random variables and I showed you that the sum of random variables is what results in a curve like this right. There is nothing much to explain there and physicists, chemists, biologists you name it people are observed this normal distribution extensively. In statistics this makes the crux of a many observations and they analyse it that is all very nicely neatly and the math of it is completely understood and it is also very nice and elegant. You may have to look at it in case you are interested ok.

And what we need to understand is we understood this; what we need to understand is this now, right. How does this combined? Let us understand this.

So, our main focus that is coming next is we look at this power law networks which is the web graphs and we will ask this question how did this happen and why did this come here, what exactly made this curve appear like this. While we were actually expecting, what were we expected? You are expecting something like this right; something like this we were expecting, we ended up getting something like this ok. We will see more of it in the next lecture.