Wireless Communications

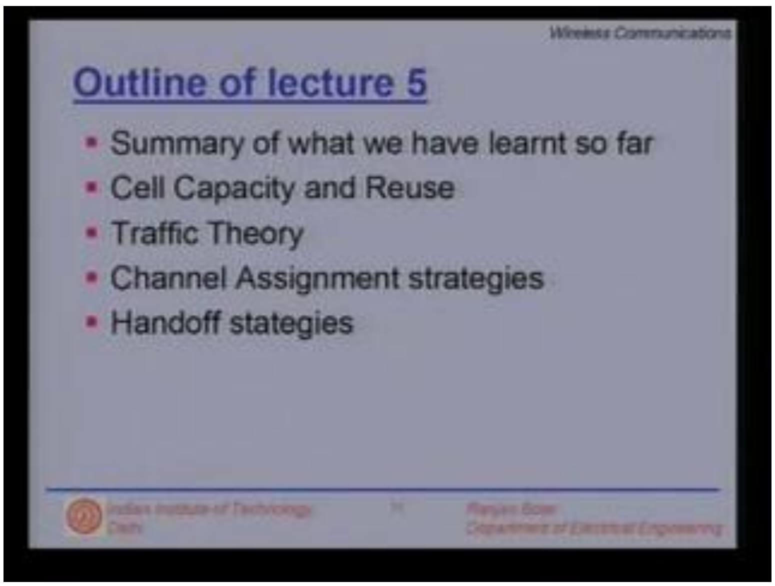
Dr. Ranjan Bose
Department of Electrical Engineering
Indian Institute of Technology, Delhi

Lecture No. # 5

Cell Capacity and Reuse

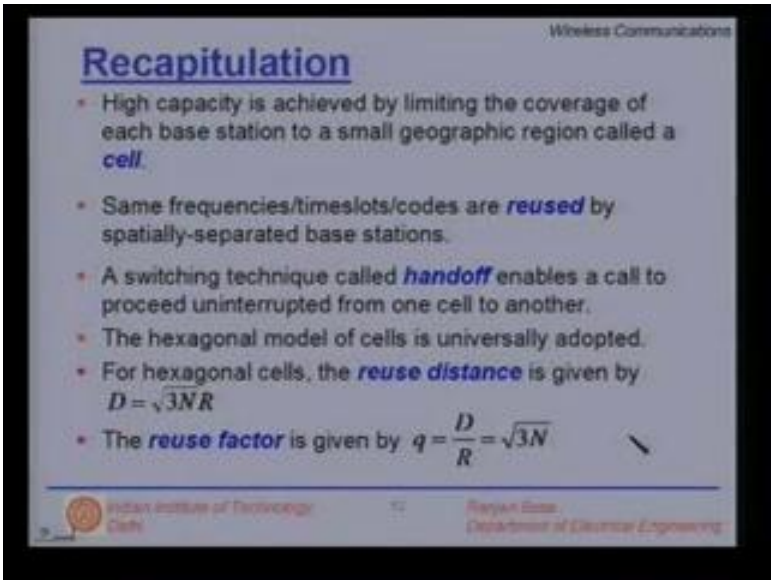
We'll look at some the interesting features of wireless communications networks especially from the cellular communications point of view. First, the outline of today's lecture.

(Refer Slide Time: 00:01:29 min)



Firstly, we'll summarize what we have learnt in the last class. We will look at an interesting relationship between cell capacity and the reuse factor. the other interesting thing that we'll learn today is traffic theory, how many calls can be handled, how many users can we support, how many calls will get blocked, how many calls will get delayed etc. all these are questions which will ultimately relate to the quality of service then we'll talk about channel assignment strategies and finally handoff strategies. So this is the brief outline of today's topic. Let us first start with summarizing what we did in the last class.

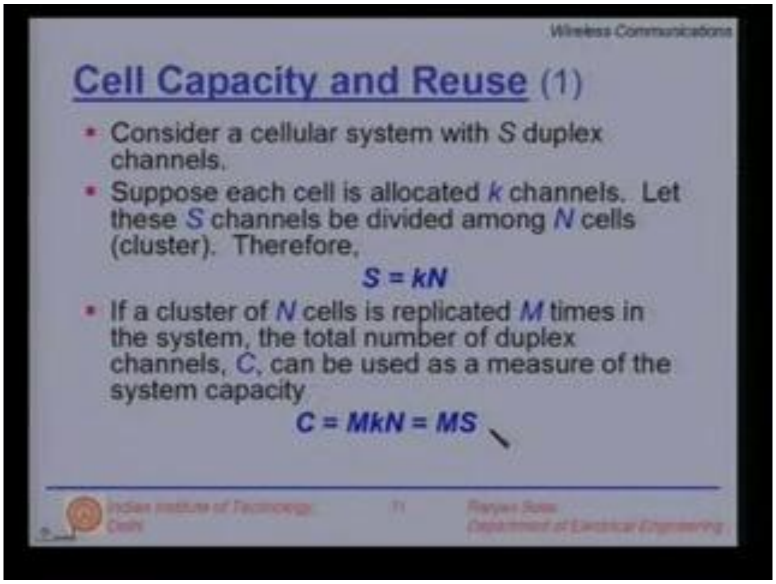
(Refer Slide Time: 00:02:22 min)



We learnt that high capacity is achieved by limiting the coverage of each base station into smaller geographical regions called cells. This forms the essence of why within the limited frequency spectrum we can still provide service to hundreds of thousands of users. The same frequency or the time slots or the codes are reused by spatially separating the base stations. We ensured that the separation is such that the interference from these co-channel cells does not exceed beyond the acceptable limit. A switching technique called handoff enables a call to proceed uninterrupted from one cell to another. Today at the end of the talk we'll look at some handoff strategies. We also learnt that the hexagonal model for cells even though universally accepted, is only of academic interest. In reality, the cells are not hexagonal. They are not circular.

They are irregular in shape. Several things determine what is the shape & size of the cell. However, when we do our basic calculations for hexagonal cells, the reuse distance is given by D = under root 3 N times R where N signifies the number of cells in a cluster. R signifies the cell radius. It is also equal to one side of the hexagonal. We also learnt last time that the reuse factor is given by D over R. it is in fact the normalized reuse distance. It normalized with respect to the cell radius R and it is under root 3 N. please remember this capital N is a design parameter. We can increase the value of N and it will lead to reduced interference. However it will also reduce our capacity. So it appears we need a handle on how to link this value of N to actual capacity of the cell and the capacity of the entire system. By capacity we mean the number of users then can actually be supported in the whole geographic region.

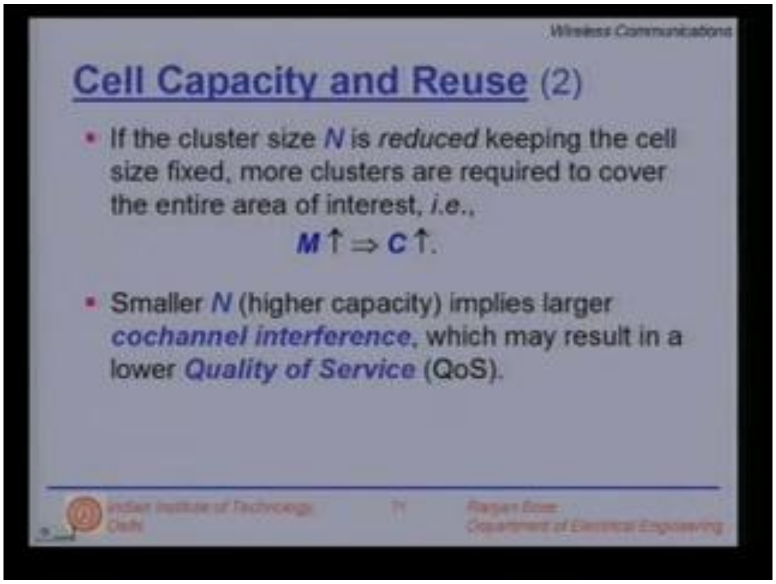
(Refer Slide Time: 00:05:11 min)



So in this slide let us have a look at the relationship between cell capacity and the reuse.1et us consider a cellular system with S duplex channels. Suppose each cell is allocated k channels and there are total N cells in the cluster, remember in the last class we saw that this value of N is an integer can be 3, 4, 7, 12 and so on and so forth. It can be 6 because of the hexagonal cell shapes that we have chosen for analysis.

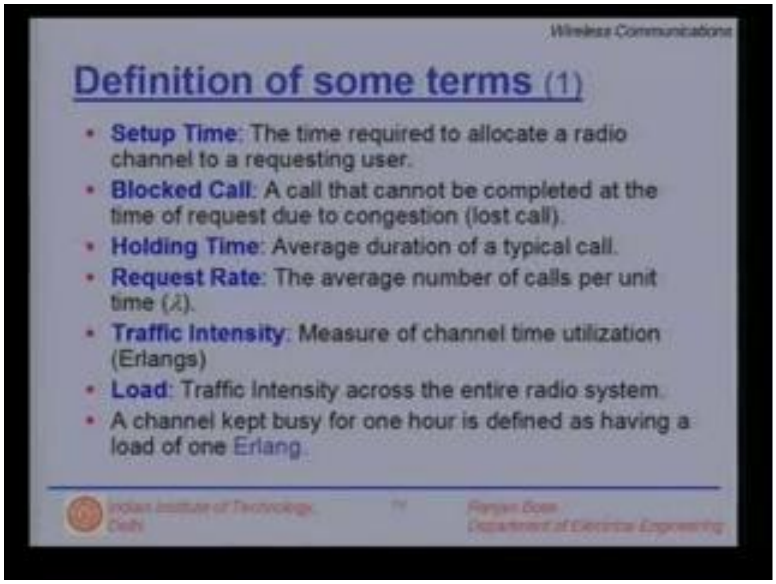
Whatever be the value of N, the total number of channels S is nothing but the number of channels in each cell times the N, number of cells per cluster. if a cluster N is now replicated M times throughout the system for example, if we have a seven cluster cell then over Delhi we can repeat it probably 20 times. So we can repeat it M times over the entire region of interest. The total number of duplex channels available, C is nothing but M times kN. M is the number of time we repeat the cluster, k is the channels per cell and N is the number of cells per cluster. Interestingly k into N is S which is a constant pre allocated number of duplex channel in the whole geographic region. S depends on the total bandwidth given to the operator. It is a constant. So we have an interesting relationship C = M time S which is a constant.

(Refer Slide Time: 00:07:11 min)



What do we conclude from this? If the cluster size N is reduced that we have fewer and fewer number of cells in the cluster, then clearly we need more number of such clusters to cover the entire geographical region. So N must go up. If N goes up, the capacity C goes up because C is equal to M times S. it could be intuitive also. But it is very a clear depiction here that "yes. If I need to support more number of users tomorrow, I have to reduce my cluster size". But of course there is no free lunch. There is a price to be paid for this. The price should come in terms of a co-channel interference level. If the cluster size is reduced, my frequency reuse distance which is normalized to root 3N goes down. Then my co-channel cell using the same frequency comes closer and my interference level must go up. Let us now look at some terms which we will use to understand how to characterize a traffic how to say how many calls will go through or how many calls will get blocked, what do we mean by trunking and how do we use certain kinds of offered loads.

(Refer Slide Time: 00:08:49 min)



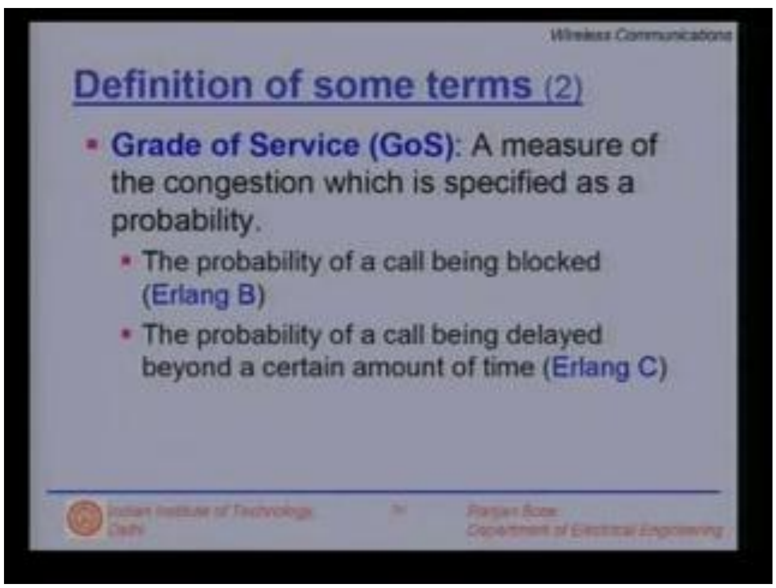
First of all, the set up time is defined as the time required to allocate a radio channel to a requesting user. Please remember there is a finite time that will elapse from the start of a request to you being actually granted the call. In fact, what is normally done is the total number of users is much larger than the total available channels at any time. We believe that not everybody is going to talk at the same time. This is the faith we have and based on this faith, we always keep the number of users much larger than the total number of channels available. However, if someone request for a call to be setup, we give one of the available channel and for the entire duration of the time the user gets to use the channel. After the call is over, the channel is returned to the channel pool. So this is the notion of trunking.

You have a trunk which has several channels. It’s a pool of channels which the users can request for and use if available. It also means that sometimes when you make a request you may not be granted a channel. You may be blocked or you may have to wait. So they are two parameters. You may have a probability of blockage and a certain duration of wait before you actually get to use the channel. both this things are probabilistic and sooner or later, we will have to come up with formulae to estimate the time you have to wait or what is the probability of blocking that will take place. So we have already defined set up time. The other thing that we need to know is what is a blocked call. A call that cannot be completed at the time of request due to congestion in literature it is sometimes called a lost call. Lost call is lost revenue. I would not like to have many blocked calls.

In fact, close to the New Year’s even if you will find most of the mobile phones are facing the problem of blocked calls, it’s because too many people are trying to make the call. Suppose you are lucky and you actually get into a conversation, you will use the channel for certain duration of time or some people talk longer than others. So there is a notion of average duration of a typical call. That is called the holding time. This holding time can be regulated by increasing or decreasing the cost per minute. So the mobile companies can actually play around with this holding time by changing the tariff. Today, usually the tariff is going down because of competition. So holding time is actually increasing.

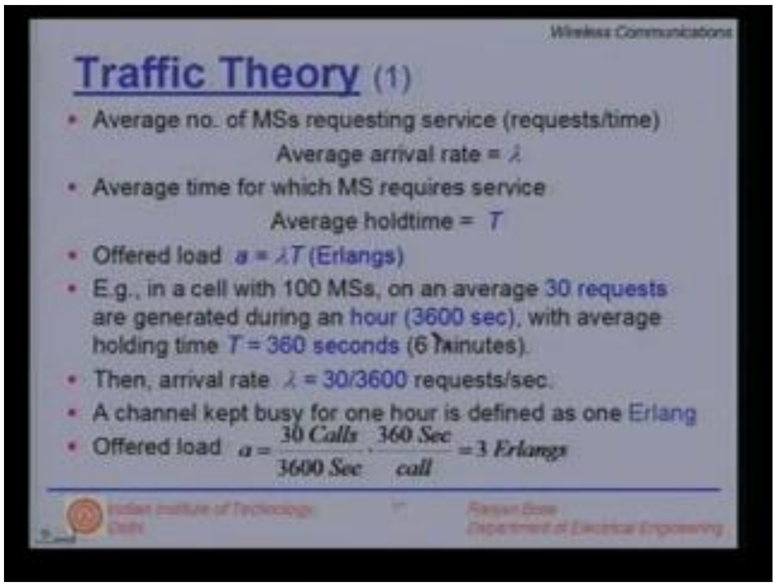
The other interesting thing that we cannot do much about is the request rate. Please remember we are in a very dynamic system. We do not know how many people are going to make that call. So the average number of calls per unit time usually represented by lambda is the request rate. So the units could be number of calls per hour. Typically in an office environment, it could be 5 calls an hour. From home environment, it could be 2 calls an hour. If it is a call center, there may be 35 calls an hour or more. So lambda changes from circumstances to circumstances. Traffic intensity is defined as the measure of channel utilization time. It is measured in Erlangs. Very soon we will define what is 1 Erlang. The load is actually the traffic intensity across the entire radio system. Please remember all of these definitions are in the average sense. A channel kept busy for 1 hour is defined as having a load of one Erlang. So that is a definition of one Erlang. We can measure the load in terms of Erlangs.

(Refer Slide Time: 00:13:43 min)



We continue with definitions. Grade of service or GoS is a measure of the congestion which is specified as a probability. Please remember it should depend on many things. it should depend on the number of users who were trying to make a call and how long they hold the call once they get connected. So there are two kinds of probabilities that we might be interested in. one is the probability of a call being blocked if my service provider has designed his system for a higher probability of call blocking. Tomorrow I might switch over to the competitor. However the design of the system for very stringent call blocking probability will require them to reserve too many channels. So we will look at an example today to see how the blocking probability translates to the number of users you can support. It’s a direct one to one relationship and directly translates to revenue. The other probability that I am interested in is the probability of a call being delayed beyond a certain amount of time. If it is beyond a certain amount of time, the user will also lose patience and probably give up trying to make the call. It is bad business strategy. So we have two kinds of charts and associated formulae. One is called the Erlang B formula and the associated Erlang B chart and the Erlang C formula and Erlang C chart which tells us directly and quickly what are the probabilities, the number of channels available and the load. We will look at this.

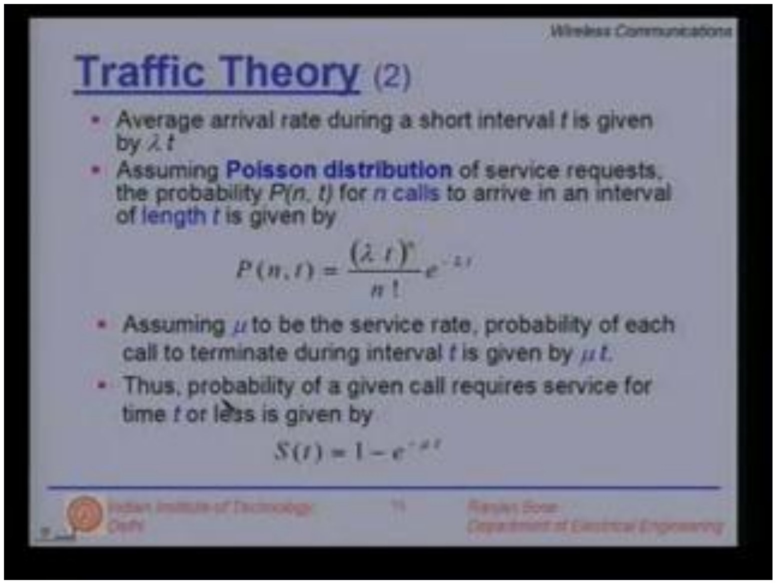
(Refer Slide Time: 00:15:42 min)



So the average number of mobile stations requesting the service could be defined as lambda. it is also called the average arrival rate. The units typically can be calls per hour. The average time for which the MS requires the service is called the average holding time T. typically, it is in minutes unless the call rates are very cheap. The offered load then is simply the product ‘a’= lambda times T and the units are Erlangs. Let us consider a small example. Let us have a cell of 100 users and an on an average, I am getting 30 requests per hour and every time a call is made, it lasts for a different interval of time. However on an average, a call lasts for 6 minutes. That is the realistic scenario that I am talking about.

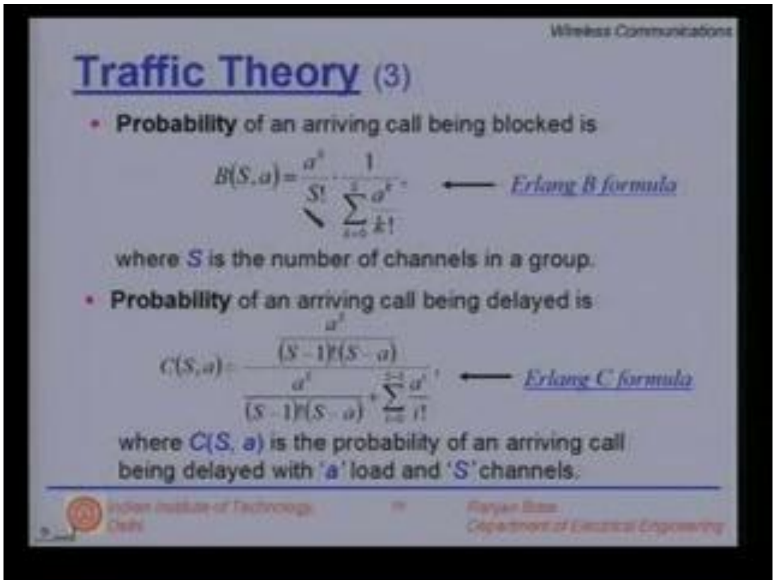
Then the arrival rate in terms of request per second is nothing but 30 divided by 3600 seconds. A channel kept busy for one hour of course is defined as one Erlang with C and the offered load which is a equal to lambda times T is nothing but 30 divided by 3600 times 60 seconds. It comes to 3 Erlangs. So I would say that the load of that system with 100 mobile stations is 3 Erlangs. Interestingly this number 100 has not come into the picture. I may have a 1000 users there but at the end of the day, I am interested in finding out how many requests they make per hour. If this is a 100 users for all teenagers and have a very busy social life, this will be more than 60 requests per hour. If these 100 mobile stations belong to a group of elderly people, it will be two request per hour. So this number 100 does not carry any meaning but this number of calls requested per hour does.

(Refer Slide Time: 00:18:07 min)



Average arrival rate during the short interval t is given by lambda t. now probabilistically speaking, Poisson distribution could be assumed as the service request and the probability P as a function of n and t for n calls to arrive in an interval of length t is simply given by P = lambda t raised to power n divided by n factorial exponential raise to the power of -lambda t. remember lambda t is of significance. it is the offered load n. it is the number of calls to arrive in a length duration of t. assuming mu to be the service rate and the probability of each call to terminate during the interval time t is given by mu times t. It comes from the basic probability theory. Thus the probability of a given call requires service for time t or less is simply given by S(t)=1- e raise to power - mu t. so what do we do with these? We can probabilistically predict what the probability is of a given call requiring service for a certain time t or less. This t will come from other system design parameters.

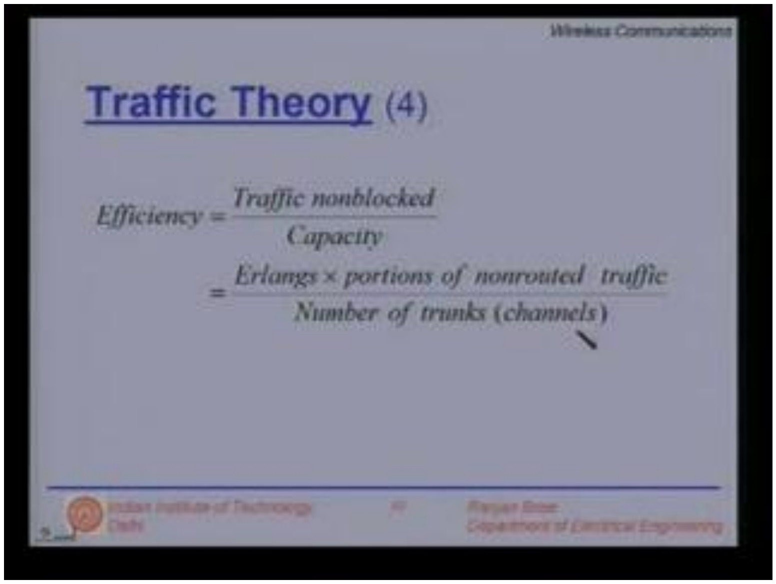
(Refer Slide Time: 00:19:32 min)



Continuing with the traffic theory, the probability of an arriving call being blocked is given by B (S,a) is the number of channels in the group. Soon we will look at a an example to see how we use this formula. But right now please remember ‘a’ is the offered load. a raise to power S divided by S factorial multiplied by 1 over summation e raise to power k divided by k factorial. The other probability that I am going to use is the probability of an arriving call being delayed. It is given by this Erlang C formula.

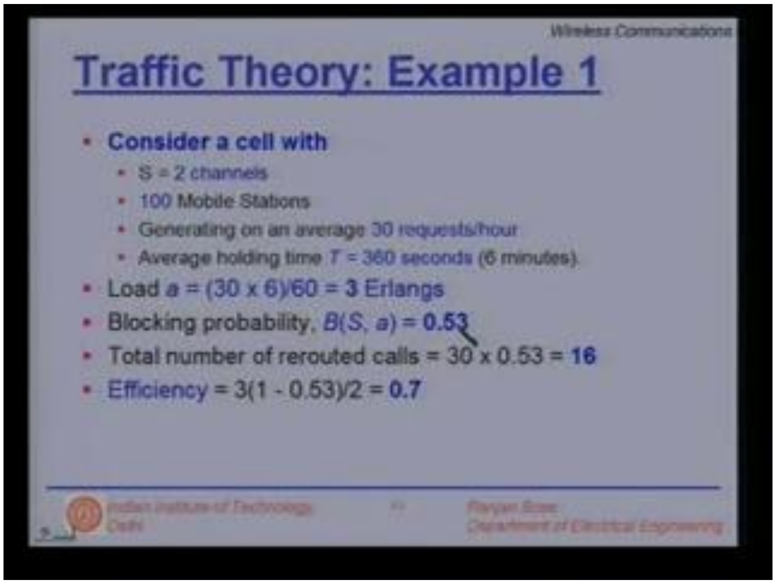
So these two Erlang B and C formulae will be used to find out. Please remember the computations can get a bit expensive. So charts are already made of the pre calculated values and they are known as the Erlang B chart which relates nothing but the probability of blocking on one axis to the number of channels on another axis and a the load on the third axis. It’s a 2D graph. So there are couple of parallel lines on the curve which relate the 3 parameters’ probability of blocking the channels and the load. Given any two, you should be able to find the third one.

(Refer Slide Time: 00:21:00 min)



Continuing with our traffic theory, we need to know something called the efficiency of the system. Efficiency is being defined as the traffic which is non-blocked and divided by the capacity. Clearly it is how many calls that I can handle divided by the theoretical limit of the number of calls that is possible to handle. It is nothing but Erlangs into the portion of non-routed traffic divided by the number of trunks. Here the word trunks signifies a cluster of channels. The trunking theory essentially realize that the number of users is much larger than the total number of channels available in the trunk. So trunk is a bunch of channels. Let’s now look at an example and see that this simple formulae can help you actually design a big system.

(Refer Slide Time: 00:22:01 min)



So here is our design space. Consider a very small cell with two channels. I have kept S =2, so that I can quickly put into that Erlangs B formula and I don't have to take in factorials of very large numbers. However I have a lot of mobile stations into this small cell. clearly it's a bad design let's find out how bad is it because I cannot have two channels supporting 100 users but remember the 100 users do not make much of an impact but how many calls they generate per hour does. so continuing from the previous example where we had 30 requests per hour and average holding time of 6 minutes, let's see under these constraints how bad or good your system design can be. the offered load 'a' can be calculated by lambda times t a simple formula if the number of request is given per hour and the call holding time is expressed in minutes, then it is request times holding time divided by 60 is units in Erlangs. it's a simpler formula. Instead of converting everything by seconds, you can still do the other way round. So 'a' is30 calls per hour times 6 minutes per call divided by 63 Erlangs.

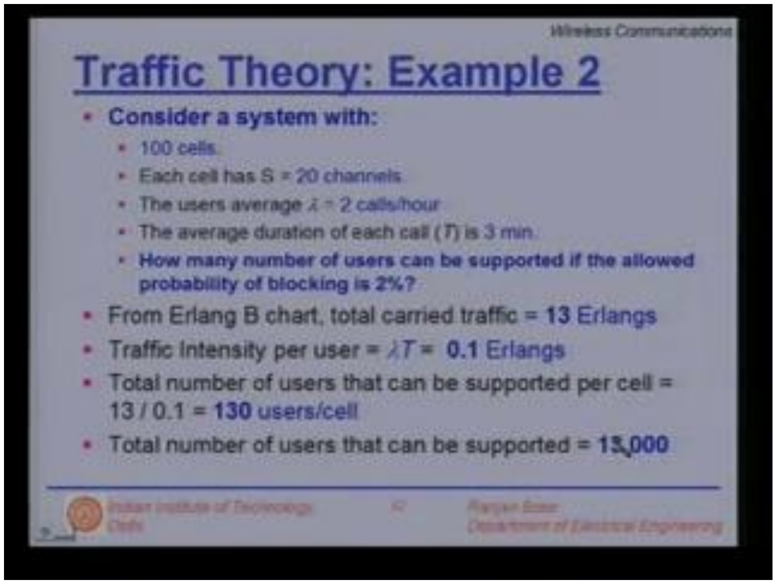
That is the traffic being generated in this small cell but a very ambitious kind of an environment with 100 users requesting30 calls per hour. So if you put these numbers where S =2 and a =3 in your Erlang B formula, you can simply calculate the blocking probability as 0.53. So more than half the calls will be blocked. It's not surprising but not too bad also. if 100 people are generating 30 calls an hour but there only two channels andthe key factor is each call lasts only 6 minutes, then probabilistically just over a little half the calls are being blocked. So whatever calls are being blocked will have to be rerouted. So the total number of rerouted calls is 30 per hour times 0.53= 16 calls per hour have to be rerouted. Efficiency given by the earlier formula is a load times 1- the blocking probability divided by the total number of channels available which is equal to 0.7. So in this sad looking channel, I still have an efficiency of 0.7. Of course this is not economically viable. Nobody will accept this kind of a service. Please remember to take this efficiency from 0.7 to 0.8, it will require you to have some more additional channels. To take it from 0.8 to 0.9, it will require you to do a much more effort. For taking it from 0.9 to 0.99, it will take you a lot more effort. So taking to that closer to 100% efficiency requires a lot of effort.

Conversation between student and professor: The question being asked is: what do we mean by rerouting a call? Suppose I try to make a call and I get the recorded message “please try after sometime”, it means my call is blocked. Now if I want to really make that call, I will actually wait for some time and redial. Again two things can happen. I might get blocked again or my call will go through. If my call goes through, my call was rerouted. In the sense that even though the first time it failed, I will have to, as a system provider route that call sometime. So if in the previous example,16 calls were blocked and these calls were really important to the users, they will make that call again and that traffic has to be handled at sometime whether I borrow some channels from the neighboring cells or I pass on this call to my competitor is my problem. So what we will do is at the later part of today’s talk, we will see that sometimes you may allow one base station to borrow certain voice channels from the neighboring base stations.

Conversation between student and professor: the question being asked is: if I borrow some channels on a temporary basis, does it have anything to do with interference? The answer is no. interference right now has everything to do with how far your co-channel cell is right and the planning has been done for that. So if I borrow some channels from the neighboring cells, I am going to also borrow certain frequencies. What is the frequency reused distance from this specially borrowed channels? So that might translate into some added interference.

Conversation between student and professor: The question being asked is: does the total load include the rerouted call? The answer is yes. This is the concept of offered load. That is, my system which consist of the users is offering a certain amount of load. In the previous example, that was fixed as 3 Erlangs. No matter what I do, the nature of the subscribers are such that they are together offering three Erlangs of load. Now whether I can handle it or not is a different story. Here the blocking probability is 0.53. Typically if I am to stay in business I must offer something like 2 %or less blocking probability. Let us look at some more examples to see how this blocking probability will actually affect the number of users I can support and that translates to the traffic.

(Refer Slide Time: 00:28:45 min)

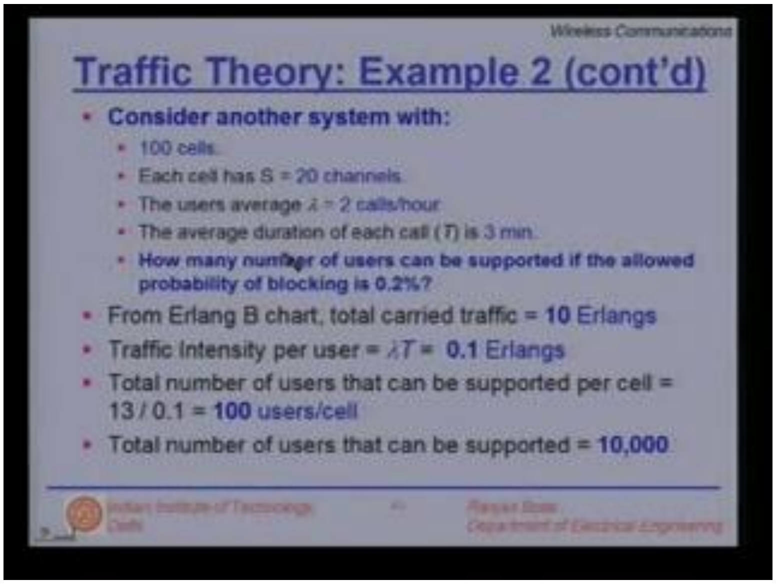


So let us look at another example. This time I am looking at a slightly larger system. The system has 100 cells and this time, each cell has 20 channels. The users are averaging 2 calls per hour and every time we make a call, they only make it for an average duration of 3 minutes. So I have changed some parameters here. It’s a more realistic situation. The questions I would like to know is: how many number of users can be supported if the allowed probability of blocking is 2 %? That is, on an average, out of every 100 calls that I have made, to get blocked I am allowing so much. Then how many users can be supported? Clearly there has to be a link. So from the Erlang B formula or a chart, the total carried traffic is 13 Erlangs. You can use these parameters to find out from that formula what the load is. So the formula gives the blocking probability. Here you have been given the blocking probability.

You need to do a reverse calculation. So you need a chart. I have already plotted all the values. I look up the chart and find out that for a total carried traffic of 13 Erlangs, the probability of blocking is 2 % provided the number of channels is20. So what are the 3 parameters I have used? 20 as the number of channels, blocking probability 2 % which gives me the third dimension 13 Erlangs. But per user traffic intensity is only lambda times t which is equal to 0.1 Erlangs. We know how to calculate this too. The total number of users that can be supported in one particular cell is nothing but 13, the total carried traffic divided by 0.1 which is equal to 130 users per cell on an average. So ideally I can handle about a 130 people per cell who are on an average making 2 calls an hour and each time they make a call, they talk for an average of 3 minutes.

My cell is not just to one but a 100 cells to cover the geographic region. So the total number of users in my entire system is 13000. It’s not very good because in a city like Delhi, this is peanuts. I need to do much better than this. I just cannot give service to 13000 people.

(Refer Slide Time: 00:31:49 min)

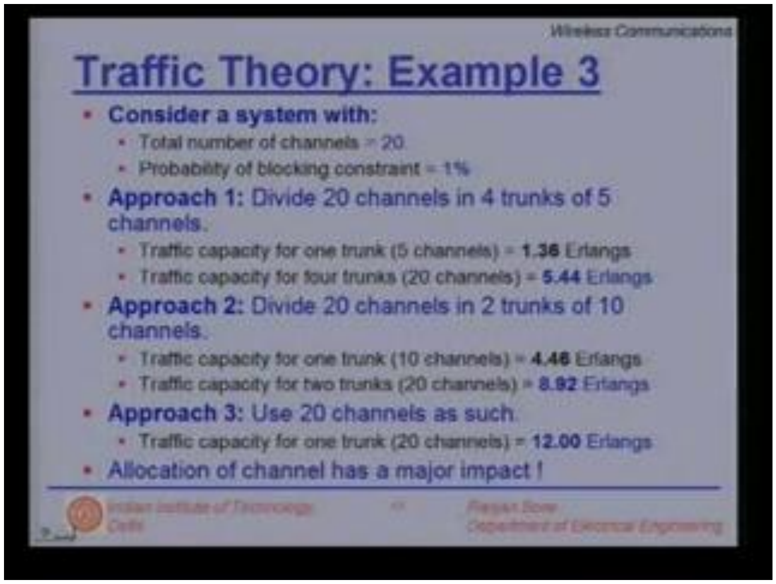


Let us continue this example and see how things will change if my competition says "look, I will give you a much better quality of service". in this case, it's called the grade of service. What does he advertise? He says, "Forget about a blocking probability of 2 %. What if I promise you a blocking probability of 0.2 %?" 2 calls out of 1000 will get blocked. so I am not annoyed each time I try to make a call during my peak office time and I say "please call back after sometime." so let's look at the competitors. He works again with 100 cells to compare apples with apples. he again has 20 channels per cell. The users are the same. They are making 2 calls an hour on an averagelasting3 minutes.

But the question now is how many number of users can be supported if the allowed probability of blocking is only 0.2 %. so intuitively what do you think? What should happen? Should I be able to support more number of users or less? Fewer numbers of users can be supported. I am talking about an order of magnitude difference in the blocking probability going from 2 %to 0.2%. Earlier I could support 13000 users. Now will it be 6000,8000 or 10000? Let's have a look at it. The answer again has to come from the Erlang B chart. So from the Erlang B chart, this time the total carried traffic is 10 Erlangs. Going from 2% blocking probability to 0.2% blocking probability, I have moved from a load of 13 to 10 only. This means that with the same traffic intensity per user as calculated before as 0.1 Erlangs, the total number of users that can be supported per cell is now 10 divided by 0.1which is equal to a 100 users only. 13 was the previous number. it should be 10. So only a 100 users per cell and over all, I can have for 100 cells, a total of 10000 users. So by going from 2%blocking probability to 0.2%, we have cut out 3000 users in my whole system.

So the question is: can I get additional users by promising to offer them 0.2%blocking probability as opposed to retaining my 2% blocking probability and supporting 13000 users? So it's a business decision at this point of time. So where do you draw the line?

(Refer Slide Time: 00:35:01 min)

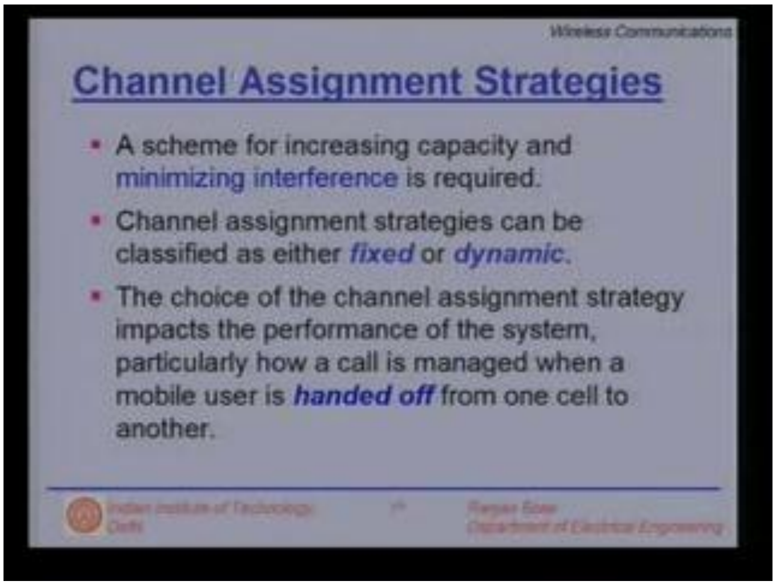


Let us now look at another example. Consider a system with a total number of channels equal to 20 and the probability of blocking constraint as 1%. So on an average 1 call in100 gets blocked. What I would like to know is that how can I organize this number of channels? That is, what is the good way to trunk? How to group the channels? if I have a 20 channels available to me, I can keep all the 20 channels as it is or I can divide the 20 channels into 2groups of 10 each called 2 trunks of 10 each or I can take these20 available channels and divide it into 4 trunks of5 each. Each trunk should provide me certain offered load. So let me rephrase the question. How should I use my 20 channels.20 is an example. if it is 200, it will make a lot of difference. So how do I use my 20 channels? Should I use it as it is? Should I group it into 2 sub sections and try to transport my traffic over these 2 sub channels or I can breakup 20 into 4 and use these4 trunks to send the channels.

Please remember the Erlang B chart is using one parameter as the total number of channels. In the case of 4, for each channel trunk which is offering a certain amount of load carriage, it can be totaled up as multiplied into 4 which will give me the total load. So let us look at this exercise. Divide the 20 channels into 4 trunks of 5 channels each. Now from the Erlangs B chart, the traffic capacity for 1 trunk consisting of 5 channels alone is 1.36 Erlangs. But I have 4 such trunks because I have 20 channels. So 4 times 1.36 gives me 5.44 Erlangs. however if I look at approach 2, that is, divide 20 channels into 2 trunks of 10 channels each, for 1 trunk of 10 channels, we have a load 4.46. So the total capacity for 2 trunks will be 8.92. This is very interesting. I have almost a 60%increased going from 5.44 to 8.92. Let's go forward. Approach 3: don't group them. Use 20 channels as such. So the traffic capacity for 1 trunk of 20 channels is 12 Erlangs. It's way too much. It tells me it is better to make a larger pool than to sub-divide the pools.

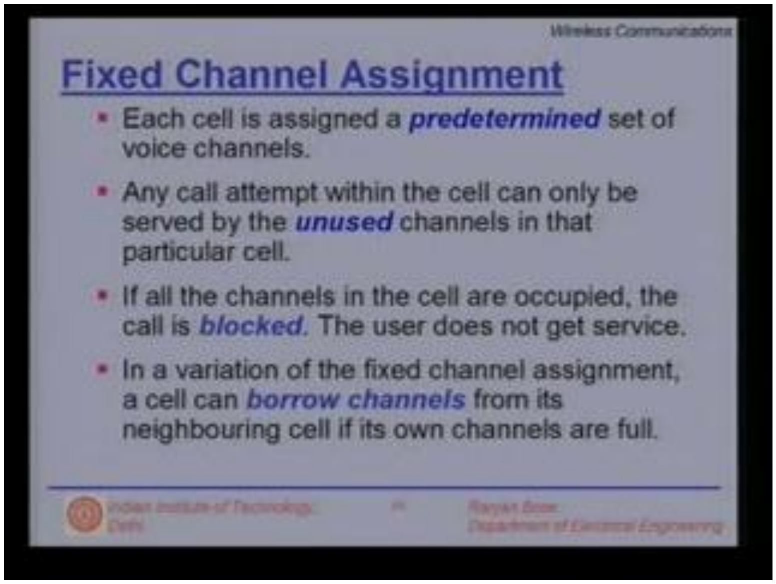
So how we group the channels, what is the trunking methodology etc will have a major impact on your system capacity. It's a very important point to be noted. So the allocation of a channel has a major impact on designing your system.

(Refer Slide Time: 00:38:45 min)



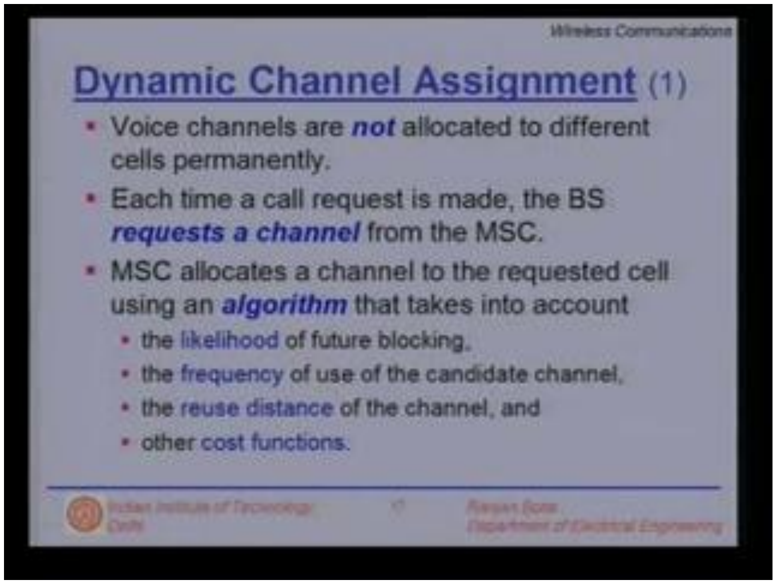
Since channel allocation is of importance, let's look at some channel assignment strategies. a scheme for increasing capacity and minimizing interference is of importance and that is what we are shooting for. Channel assignment strategies can be of two types. It can be either fixed or it can be dynamic. Fixed of course has an advantage that you plan it, you fix it and you forget it. Dynamic on the other hand requires you to be adaptive. The choice of channel assignment strategy impacts the performance of the system, particularly how a call is managed when a mobile user is handed off from one cell to another. So capacity is one perspective and the hand off is the other perspective. Later today we should also look at some hand off strategies.

(Refer Slide Time: 00:39:51 min)



So the first methodology is fixed channel assignment. Remember the previous examples we did. We tried to fix certain number of channels per cell and then we did the calculations. So let us look at the fixed channel assignment strategies. Each cell is assigned a predetermined set of voice channels. Any call attempt from within the cell can only be served by the unused channels in that particular cell. If all the channels in the cell are occupied, the call is blocked as you have seen. The user does not get service. From earlier examples, we have also seen what is the blocking probability, how can it be calculated and how can it be linked to the offered load. In a variation of the fixed channel assignment, a cell can borrow channels from its neighboring cell, if its own channels are full. This requires a little bit of intelligence which neighboring cell to borrow from and again it has some related interference issues. But still it falls under the category of fixed channel assignment.

(Refer Slide Time: 00:41:16 min)



The other face of the coin is dynamic channel assignment. Here contrary to the previous methodology voice channels are not allocated to different cells permanently. Each time a call request is made, the base station requests a channel from the mobile switching center which is the brain behind the whole thing. Remember in more complicated cellular systems there can be more than one mobile switching center. There could be a head mobile switching center which controls the sub mobile switching centers. So we can have an hierarchy. But wherever it goes, the call request is made from somebody who manages the channels. The mobile switching center - MSC allocates a channel to the requested cell using an algorithm. So it has to do some thinking. It has to have a strategy to allocate the channel. What could be the possible things it has to consider before making a channel assignment?

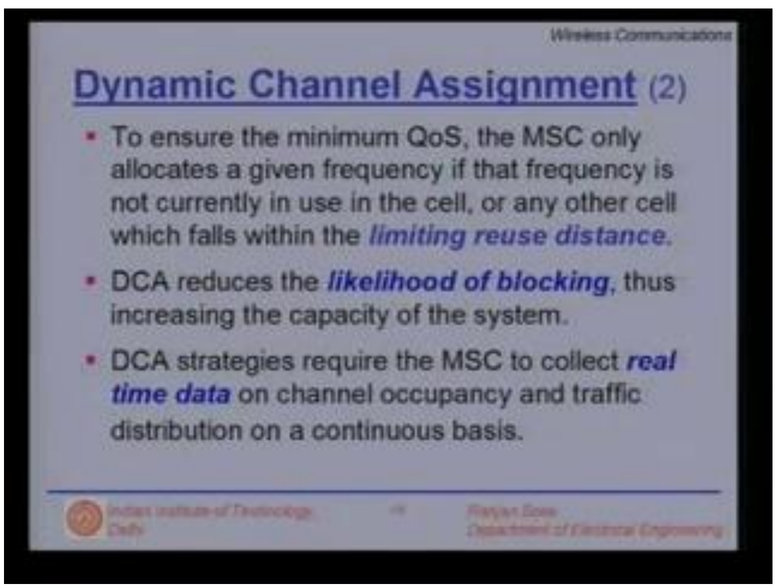
Well, it's the likelihood of future blocking. So if it has a choice of taking one extra channel from a cell in the Connaught place area where the traffic is expected to be high but currently is not, versus a cell which is in the outskirts, then it will rather borrow a channel from the cell which is in the outskirts of the city clearly. The frequency of use of the candidate channel will relate to how much service we can provide in future. About the reuse distance of the channel, the channel must be such that it does not result in the lot of interference because clearly when we borrow channels, we will cause some kind of unwanted interference because it is not as per the frequency reuse plan. So I must assign a channel from a cell so that the resulting interference is not very high and there is some other cost functions. I can have priorities built in to it, some are preferred customers and things like that.

Conversation between Student and Professor: see if my channel is being used, the request that I am making is being requested very many times, so it's a popular channel and that comes from a popular cell. So this should be taken into consideration that there will be a request very soon for this cell again. So it should have an impact on the possibility of request from that set of channels.

Student: sir, are all channels considered to be of the same type only?

Professor: Yes. Here as a first level design, we are not differentiating between any one channel with respect to another channel. However we can have variations. For example, some can be high priority channels. We can have resource reservations where certain preferred customers or who are paying a premium can always enjoy the luxury of non-blocked calls. That is, their calls will always go through. For example, suppose tomorrow, I would like to use one of these channels not as voice but for data transfer and I am a credit card company where I want none of the credit card verification to be blocked, then I will reserve a chunk of channels which will never be blocked. So those are preferred channels. So calculations for those have to be done separately. In other ways, all the channels are identical and no preference over one over the other because the bandwidth is the same, the interference issues are the same, etc. so the channels are kind of equivalent.

(Refer Slide Time: 00:46:02 min)

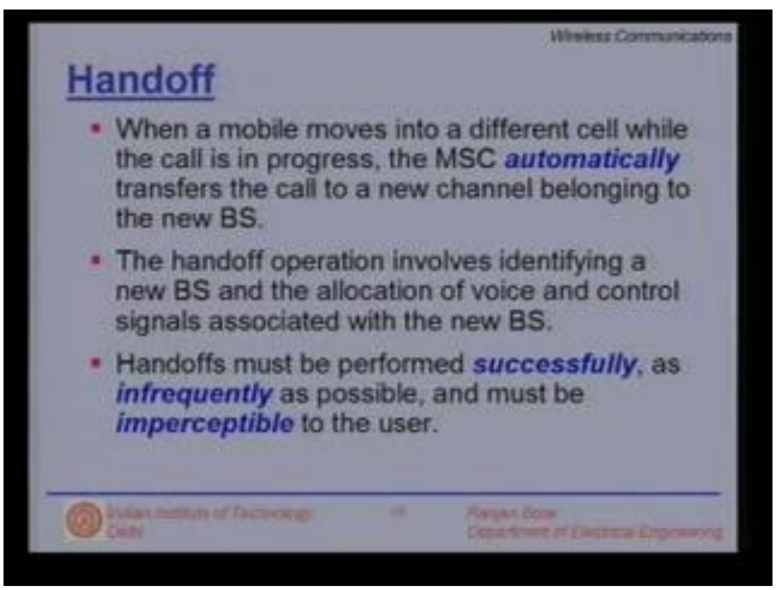


Continuing with the dynamic channel allocation, to ensure the minimum quality of service, the mobile switching center only allocates a given frequency if that frequency is not currently in use in the cell or any other cell which falls within the limiting reuse distance. Remember this reuse distance comes into the picture again because if you give a certain channel, it should not create a higher level of interference. So there is a minimum quality of service to be guaranteed. The dynamic channel allocation reduces the likelihood of blocking thus increasing the capacity of the system. Please remember we looked at the example of trunking. Fixed channel allocation is somewhat like trunking in the big scale. What we do is we group together certain number of channels and give it to different cells. Those are fixed just like total of 20 channels broken up into groups of 5. So 4 cells get 5 channels each. However in dynamic channel assignment, I have the entire pool at my disposal and as and when the calls come in, I can allocate a channel. So the borrowing can inherently happen. Clearly the capacity is much higher because the blocking probability will be much lower.

The dynamic channel allocation strategies require the mobile switching center to collect real time data on channel occupancy and traffic distribution on a continuous basis. Please note that to make any of these algorithms effective, I must have true channel data. so channel measurements have been done. It's not a big problem. All the calls which are being blocked can easily be recorded. The number of calls being generated can also be recorded. You can calculate lambda. You can calculate T, the call holding time by monitoring on an average, how many calls have been made and held during the call. Next we come to the issue of handoff. Please remember handoff has to be kept in mind as long as we talking about mobile cellular services. Handoff so far we have you seen can happen because of two things.

Firstly you move from one cell to another cell while you are still on the call and the second thing is you may be required to handoff one channel and go on to the next channel. So there are two kinds of handoff that takes place. Let's look at the handoff required when you move from one base station to the other base station. So when a mobile moves into different cell while the call is in progress, the mobile switching center must automatically transfer the call to a new channel belonging to the new base station.

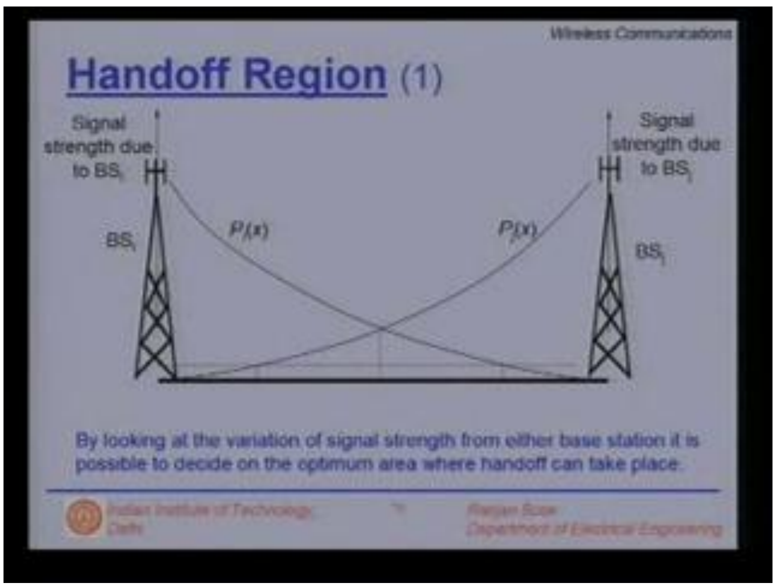
(Refer Slide Time: 00:49:02 min)



It’s a new area, new base station and a new channel. The handoff operation involves identifying a new base station and the allocation of voice and control signals associated with the new base station. All this must be done seamlessly. The handoffs must also be performed successfully as infrequently as possible, and must be imperceptible to the user. Why should it be infrequently?

This is because each time I do a handoff, I use overheads .in fact, there are overhead channels used for handoffs alone. so it requires computation on the part of mobile switching center and if too many mobile stations are trying to handoff at the same time, the mobile switching center might actually get bogged down by the number of computations it has to do because remember there is an algorithm to be followed while doing handoff. So it should be done as infrequently as possible. Let’s look at the handoff region pictorially.

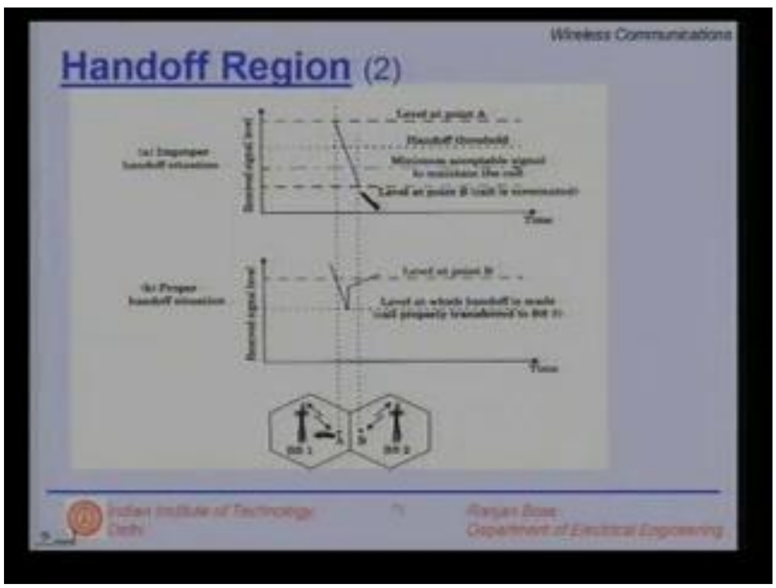
(Refer Slide Time: 00:50:25 min)



So what we intend to say in the slide is by looking at the variation of signal strength from either base station, it is possible to decide on the optimum area in which the handoff can take place. So the handoff region is not a discrete line that you draw. You belong to base station 1 if you are on the left side of this line and base station 2 when you are on the right side of this line. In fact you have a region where you can handoff. Now the size of the region is important with respect to the velocity of a travel. If you are travelling too fast, you have very little time to handoff and so your call may get dropped during handoff.

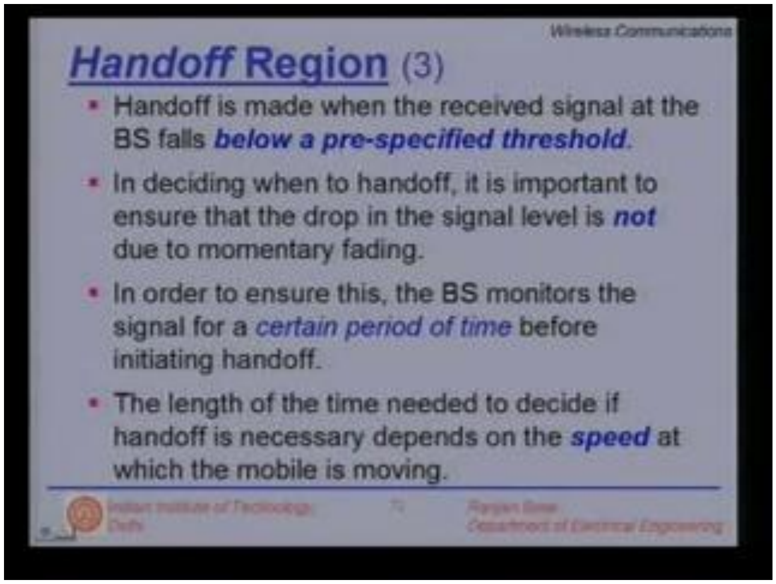
We do not want this. So let us look at it. Suppose we have base station I and we have another base station J. they are located at a certain distance apart. Now what we would like to do is have a mobile station which hopefully is now communicating with base station I. but it’s travelling towards base station J. during the process of travel, the signal strength from base station I starts dropping and as it goes towards the base station J, it goes below an acceptable region. if you reverse the situation and make sure that the mobile station starts from base station J but this time moves towards base station I. again the signal must drop as it moves away from base station J but clearly goes closer to the base station I. so we have a parallel process where signal strength slowly goes down as you move away. Now these curves are very smooth because we are not taking into consideration the fading effects. Fading will be taught at the later part of this lecture series. Clearly we have two plots where power levels are going gradually down as you move away from base station I and on the other hand, going down as you move away from base station J. I would like to know what is the region where I can handoff? So for that I need to know what is the minimum sensitivity of my receiver or even beyond that, what is the threshold till which I can still make a good call. So it is nothing but a line here because on the y axis I have plotted the signal strength. So the horizontal line intersects the two curves at point 1 and point 2. Anywhere in between, my mobile stations can handoff. But if it moves fast, it might cross this region in 3 ms. if it is a pedestrian he may just be there for 30 minutes. so my handoff is much simpler for a pedestrian traffic than for the vehicular traffic.

(Refer Slide Time: 00:53:57 min)



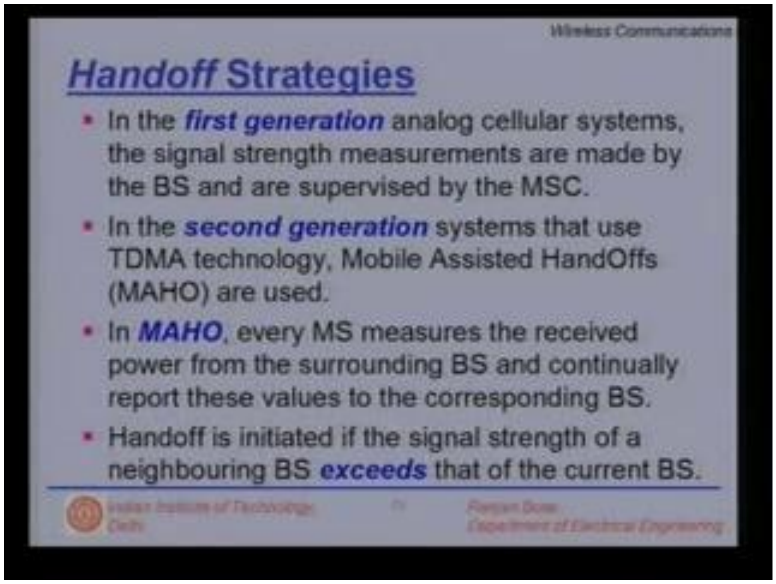
In this slide I have try to show two circumstances. One where the handoff occur successfully in the lower curve and one when it fails. So let’s look at the lower curve first. Remember there are two regions base station 1 and base station 2 and I am moving from left to right. Here is a depiction of the falling of the signal strength of the first base station. but very soon it locks on to the another signal strength of base station B. so there is a sudden jump in the power signal level because I am getting a much stronger signal from base station B. so I have hooked on to the next the power level. So it’s a successful handoff. On the other hand, here I keep on dropping and unable to latch on. Maybe there was no available channel or there was some other problem and so I go below a certain level and I cannot operate beyond that.

(Refer Slide Time: 00:55:07 min)



So, handoff is made when the received signal at the base station falls below a pre-specified threshold. In deciding when to handoff, it is important to ensure that the drop in the signal level is not due to momentary fading. We will talk about fading later. In order to ensure this the base station monitors the signal for a certain period of time before initiating the handoff. So it should not be a false alarm. I should handoff when truly the signal has dropped below a certain level and not a temporary drop. The length of time needed to decide if the handoff is necessary depends on the speed at which the mobile is moving. Since it's mobile communication, I cannot put restrictions on how fast they move.

(Refer Slide Time: 00:56:01 min)

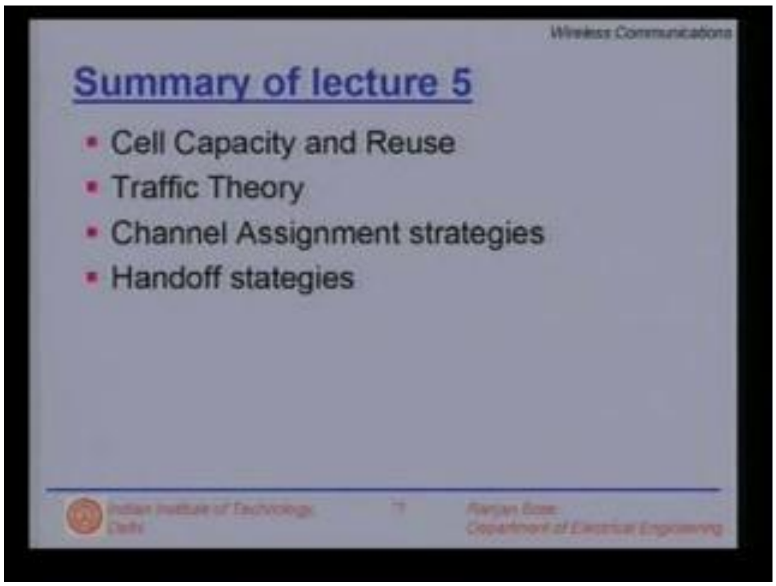


Some strategies of handoff. In the first generation analog cellular systems, the signal strength was measured by the base stations and supervised by the mobile switching center. In the second generation those which use TDMA- technology mobile assisted handoff' is used called MAHO. In MAHO, every mobile station measures the received power to the surrounding base stations and continually report these values to the corresponding base stations. Handoff is initiated if the signal strength of a neighboring base station exceeds that of the current base stations. So try to hook on to a stronger signal at any time.

(Refer Slide Time: 00:56:47 min)

In this last slide for today let's briefly look at what is the notion of soft handoff, typically associated with the CDMA systems. The code division multiple access spread spectrum cellular systems provide a unique handoff capability. Unlike channelized wireless systems that assign different radio channels during a handoff which we call "hard handoff", the spread spectrum mobile station share the same channel in every cell. Please note the difference. The same channel in every cell. The term handoff here implies that a different base station handles the radio communication task. The ability to select between the instantaneous received signals from different base stations is called a soft handoff.

(Refer Slide Time: 00:57:53 min)



At this state let us conclude today's lecture and let me briefly summarize what we have done in lecture number 5. We started off with a brief recap and then looked at the cell capacity and how the reuse affects the cell capacity. Then we briefly looked at traffic theory. We looked at Erlangs B formula& the Erlangs C formula. Then we looked at some examples as to find out how to relate the blocking probability with the offered load. Finally we looked at channel assignment strategies and some handoff strategies. Thank you for your attention.