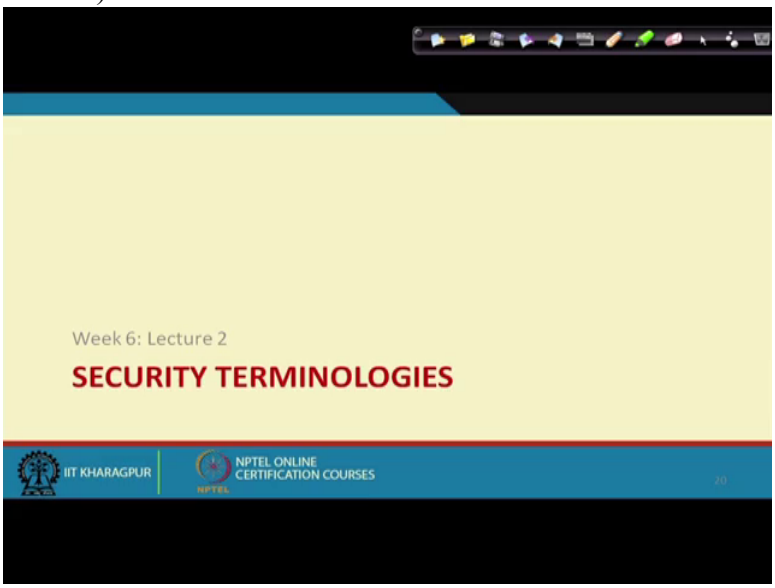
Course on E-Business
Professor Mamata Jenamani
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur
Module 06
Lecture Number 29
Security Terminologies

(Refer Slide Time 00:18)

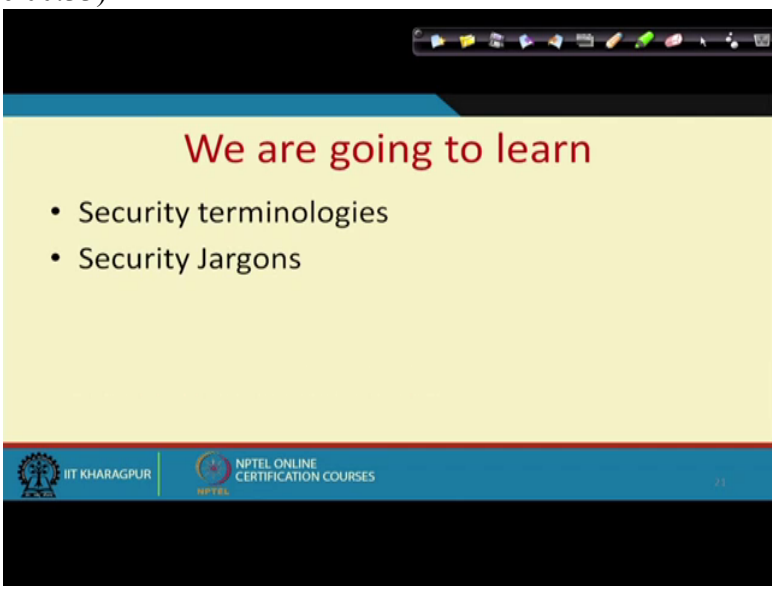
Welcome back. So last class we talked about various security categories. Now

(Refer Slide Time 00:24)



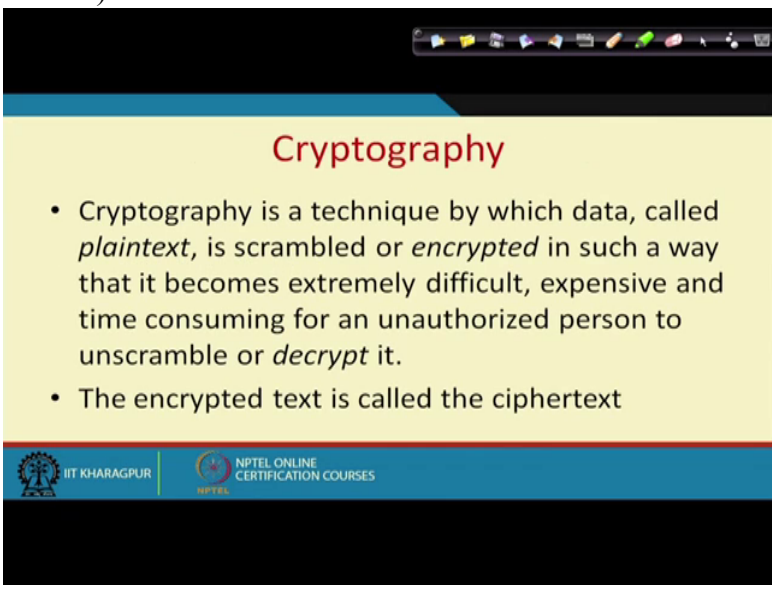
specifically we are going to look at various terminologies associated with this security mechanism.

(Refer Slide Time 00:33)



So here the first

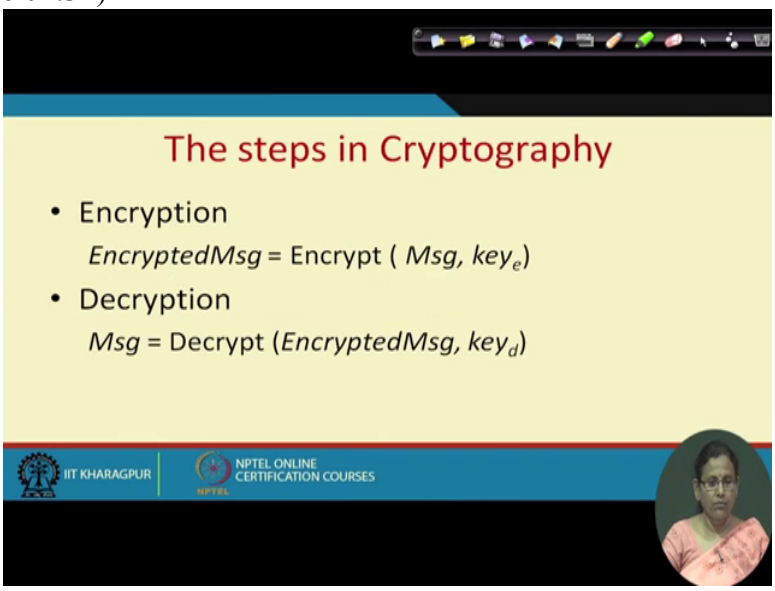
(Refer Slide Time 00:36)



terminology that we are going to talk about is cryptography. And while talking about these security categories, we learnt about, that one of the security, important security category is that the data should not be revealed to the unauthorized people. So this cryptography is the technology which helps doing that. So this cryptography is a technique by which the data called the plain text is scrambled or encrypted in such a way that it becomes extremely difficult and expensive and time consuming for an unauthorized person to scramble or decrypt it.

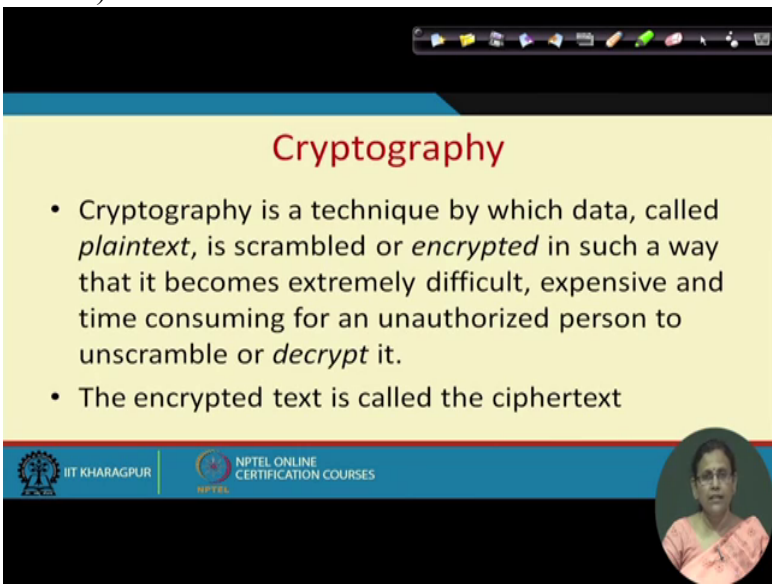
So which means now what is your plain text in this context? The application data which goes through your T C P, we were looking at this header of,

(Refer Slide Time 01:32)



in the earlier lecture

(Refer Slide Time 01:35)



we were seeing that when the application data is sent from one host to other it is broken into packets and

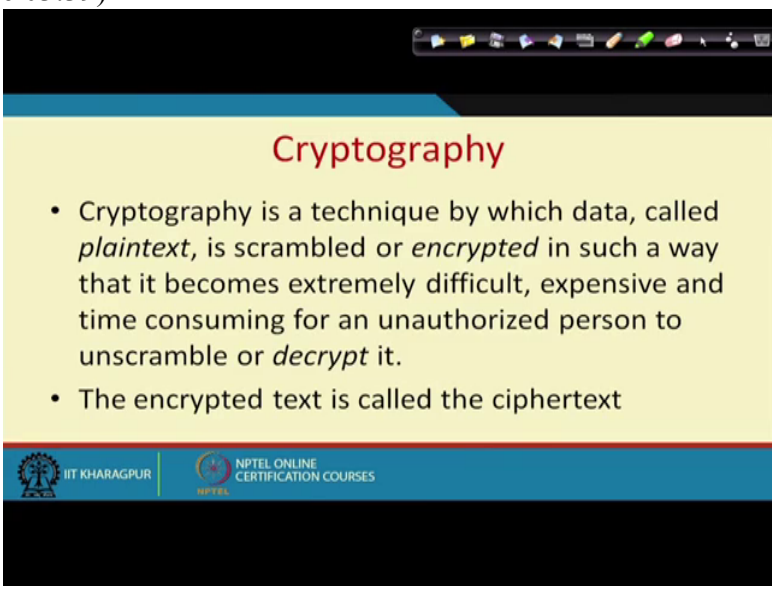
(Refer Slide Time 01:42)

headers are added. Now we are talking about, not about the header part, but about the application data part. That application data that you are sending, should now be encrypted. It has to be, some transformation has to be made to the data so that it becomes extremely difficult for somebody else even if he is, see, as I have told you, the person from, this attacker can actually take all your application data part and it can put them together according to the sequence number, T C P sequence number, and it can get the complete information.

Suppose you are sending some business document. So it is revealed. So to avoid this each of these application data part, if you scramble in some manner which is called encryption in the cryptographic language then if it is scrambled and then attacker even reads that part of the data, that part of the application data and combines it together, it is not going to get any, any, it is not going to extract the business information that you are sending somewhere, to somebody who is supposed to receive it, your competitor cannot do it.

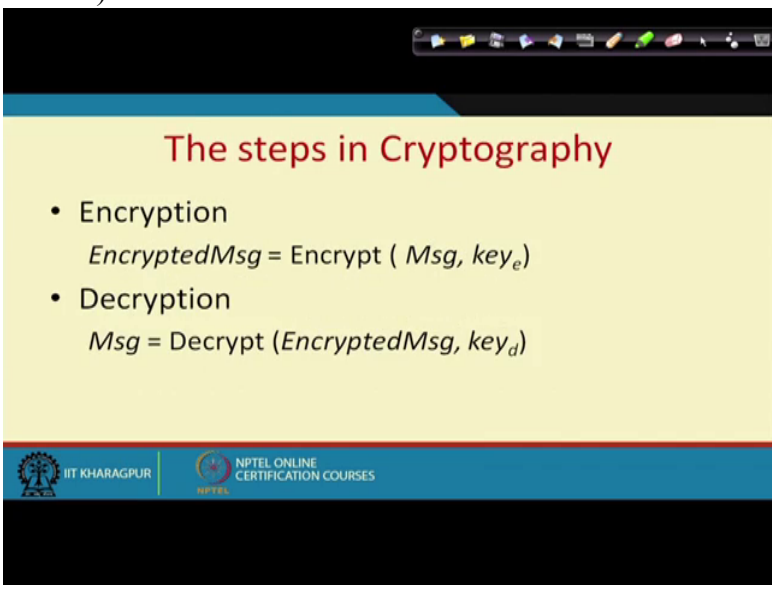
So first of all, it should be extremely difficult, or it has to be expensive. By expensive we mean the amount of effort put for decrypting that should be so expensive in terms of money as well as time, the person who is attacking loses his interest to do so, Ok. So

(Refer Slide Time 03:39)



the encrypted text is here called the cipher text.

(Refer Slide Time 03:43)

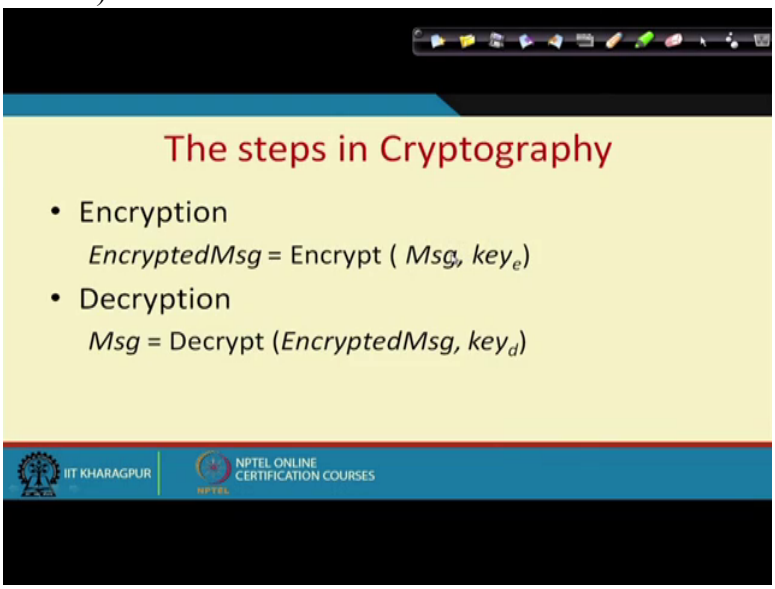


So in any cryptography, the two steps are followed. First is encryption. Second is decryption. In case of encryption, in both encryption and decryption requires some key. This message is actually that application data part. There has to be some function. I mean, they are very complex mathematical

(Refer Slide Time 04:12)

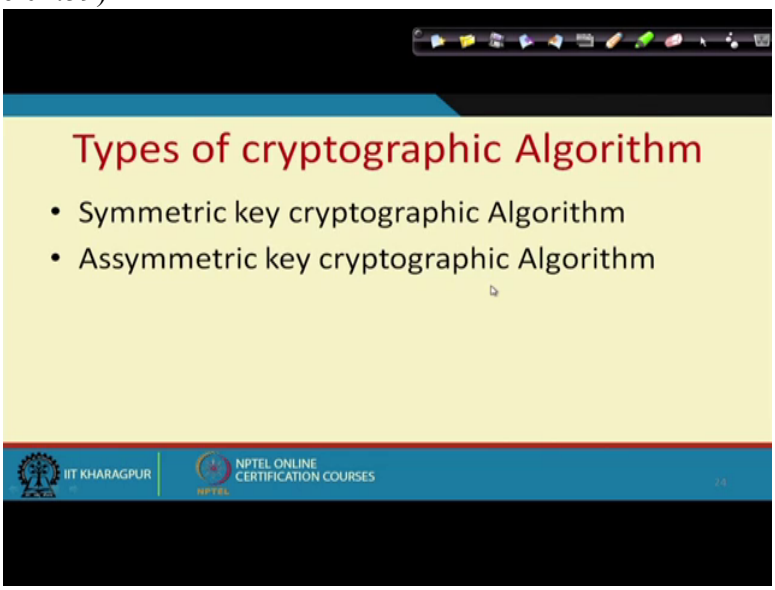
functions. These complex mathematical functions

(Refer Slide Time 04:15)



have two parameters; one is the message, the plain text that you are sending and then the key. So the function uses these two parameters to create something called the encrypted message or the cipher text.Then during the decryption process, just the reverse happens. The encryption message and the key is being used to decrypt it, to get the original message back. Now depending on the nature of this key, this first one is called encryption key, the second one is called decryption key. So in case it is a, in case

(Refer Slide Time 04:59)

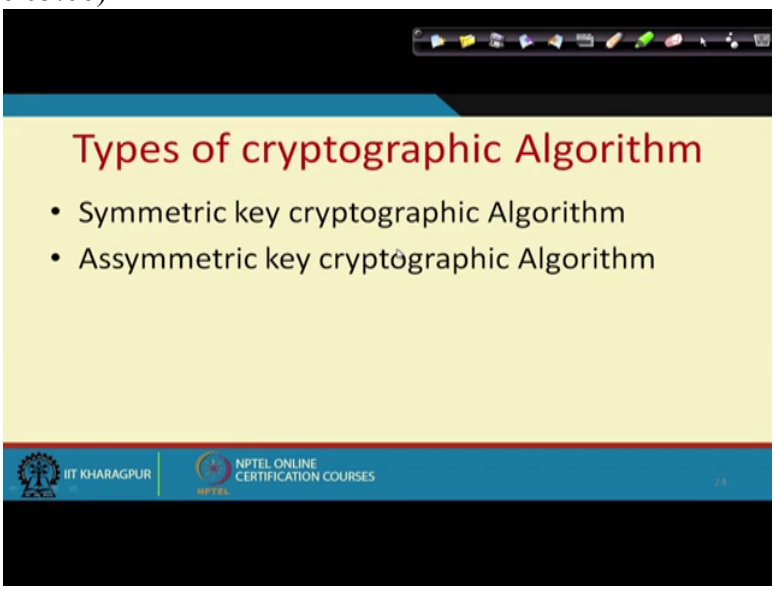


I mean depending on the nature of this key, either

(Refer Slide Time 05:03)

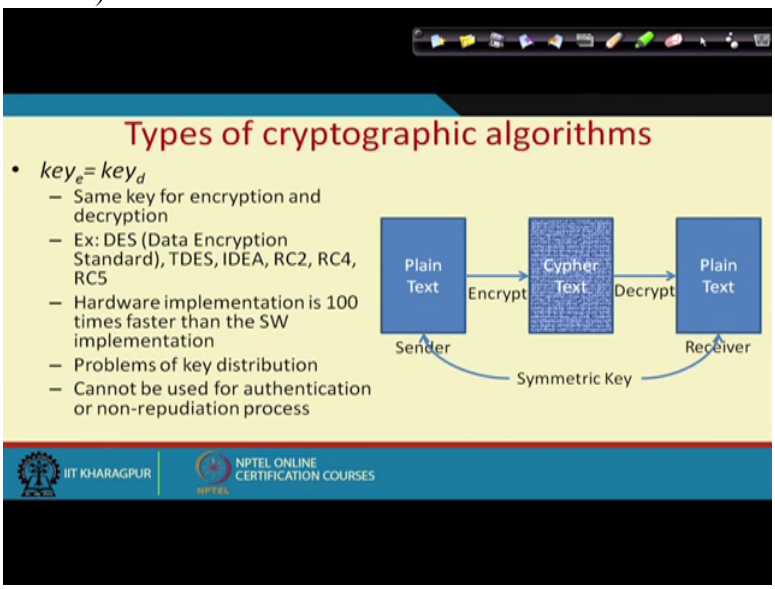
you have two different kinds of algorithms, symmetric

(Refer Slide Time 05:06)



cryptographic algorithm and asymmetric cryptographic algorithm.

(Refer Slide Time 05:11)



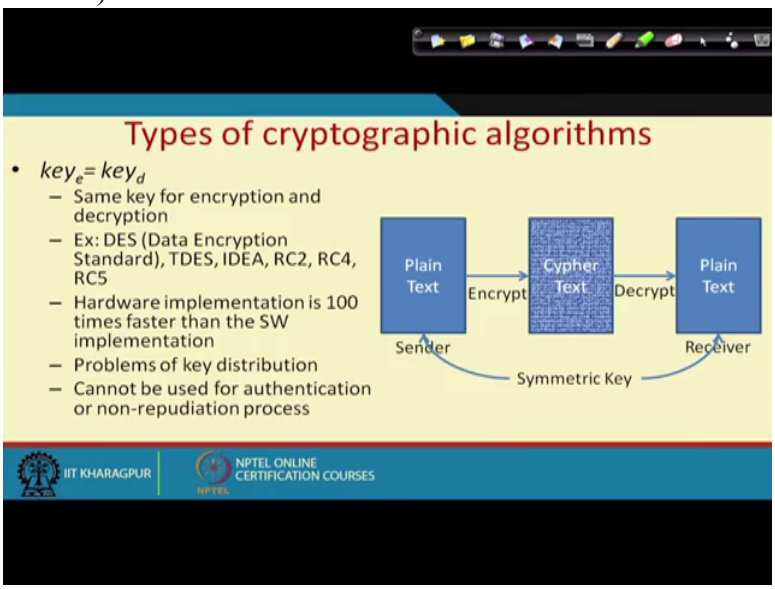
Let us look at the concept behind the symmetric key cryptographic algorithm. In this particular case, this, the encryption key, that key E is the encryption key is same as that of the decryption key. So what happens when this particular key is known to both sender as well as the receiver? So when this plain text along with this symmetric key, some encryption algorithm is applied, you generate the cipher text. That cipher text sent over the internet to the receiver who then uses this same key, same key to decrypt it. Now because both the parties use the same key, this process is called symmetric key encryption algorithm.

So there are many variants of this. First one is D E S, Data Encryption Standard, Triple Data Encryption Standard, IDEA, R C 2, R C 4, R C 5 and so on. They are basically, the main concept is same but the way, but the algorithms are different, and their strengths are different. Then such symmetric key algorithms

(Refer Slide Time 06:47)

can be either, you can have a software to carry out this cryptographic symmetric cryptographic operation, or you can use a hardware. This hardware is usually almost hundred times faster than the software implementation. Now

(Refer Slide Time 07:04)



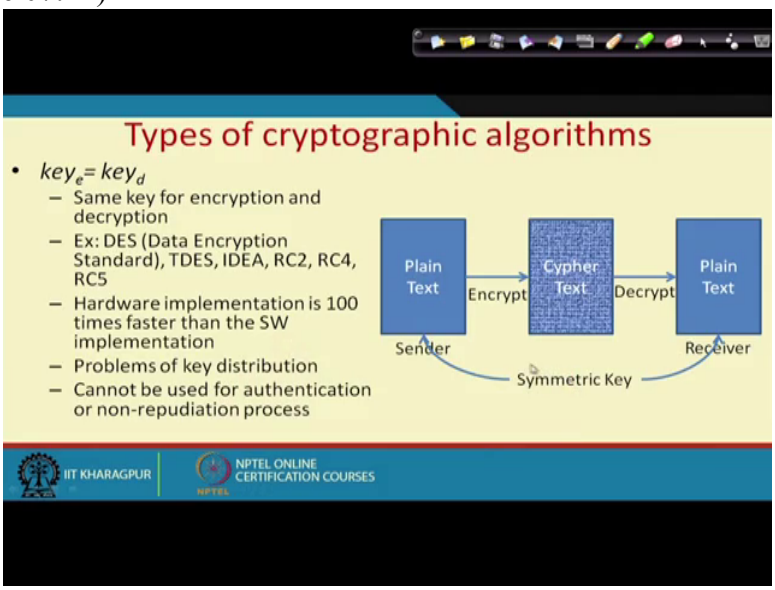
if we have such a algorithm, then what is the need for

(Refer Slide Time 07:09)

asymmetric key algorithm?

Look here though very simply we are

(Refer Slide Time 07:14)

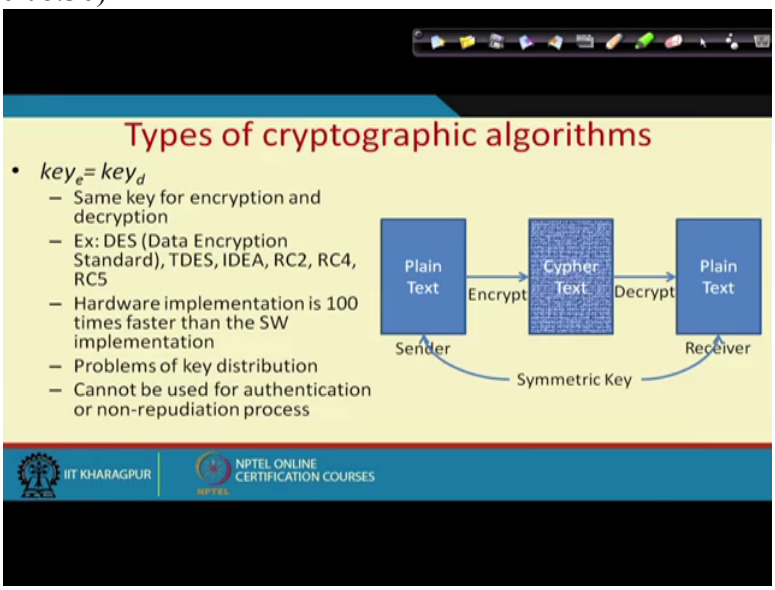


saying that the symmetry key is known to both sender and receiver, suppose you are the sender, you are in India and let's say your receiver is in United, in the U S A. So unless otherwise both of you actually know this symmetric key, how are you going to do the encryption and decryption? You might be thinking, Ok, symmetric key itself I will encrypt and I will send but that will require another symmetric key. And sending this key over this is not possible unless otherwise you have another mechanism to send this symmetric key. Look key is a

(Refer Slide Time 08:07)

very small, it is comparatively, it is extremely small compared to the actual application data. So if you have a more complex algorithm which can be used to send the symmetric key and it does not require both the parties to have the same key, then it is much better. So such algorithms are called asymmetric key

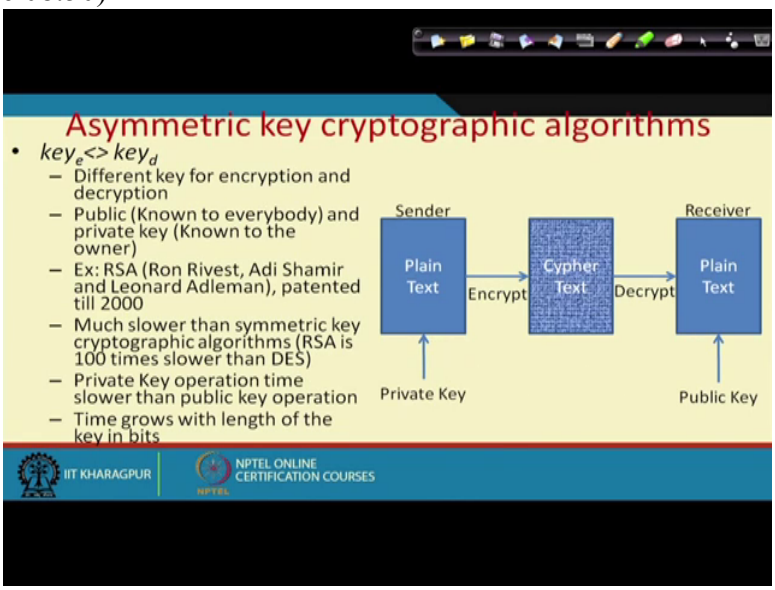
(Refer Slide Time 08:36)



cryptographic algorithms.

So in the symmetric key algorithms, the problem, the major problem is key distribution. So to resolve this key distribution process asymmetric key algorithms can be used. Now let us see what is

(Refer Slide Time 08:50)

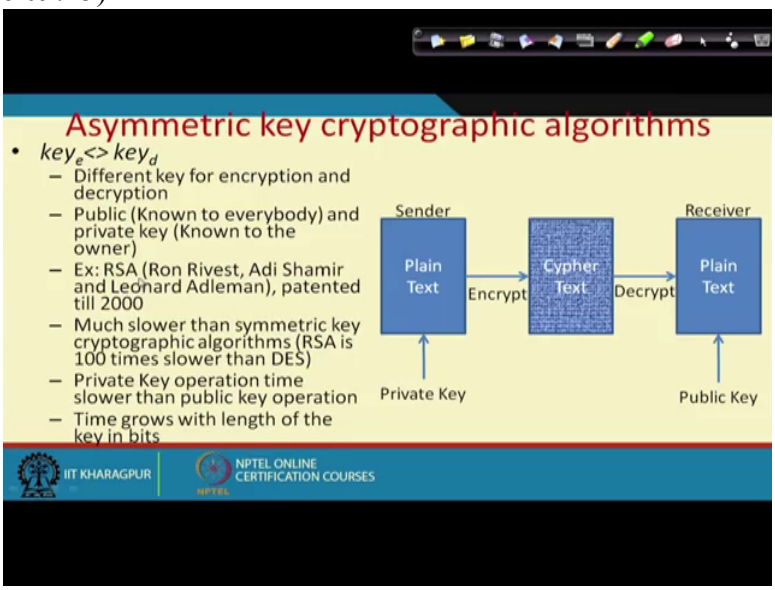


done in asymmetric key algorithm. In case of, as the name indicates, in case of asymmetric key cryptographic algorithm both the encryption keys and decryption keys are different. So the key which is known to everybody,

(Refer Slide Time 09:10)

so every entity who is participating in this asymmetric key cryptographic process needs to have two keys; one key which is private to him and another key which is known to everybody. So the one, the key which is known to him is the public key and the key which is known to him is called the private key and which is known to the, known to everybody else, it is called the public key. This particular algorithm is called,

(Refer Slide Time 09:45)

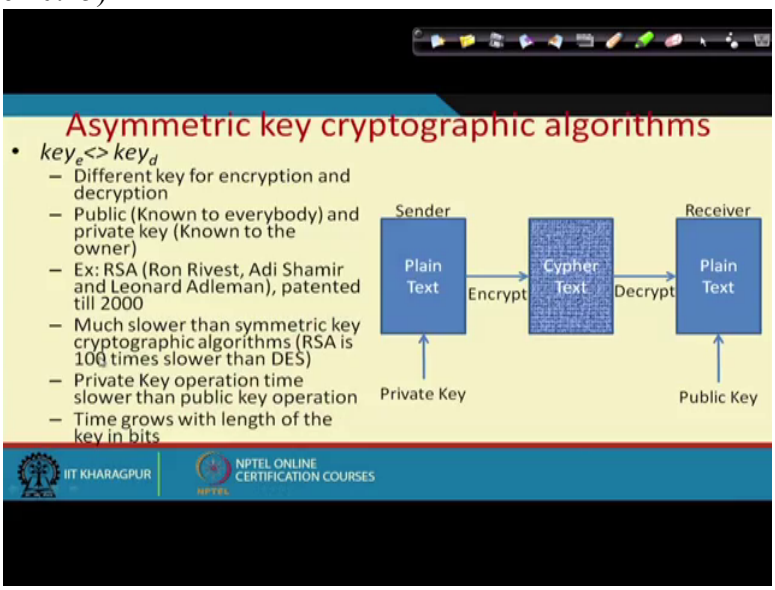


I mean there are some, not many algorithms in this category and this algorithm which is mostly used is the R S A algorithm by named after three persons who invented it and the, such algorithms again as I told you already, all these cryptographic algorithms are complex mathematical

(Refer Slide Time 10:10)

functions. So it is also a complex mathematical function and this algorithm is much, much slower than that of your corresponding symmetrical key algorithms. For example it is almost

(Refer Slide Time 10:25)



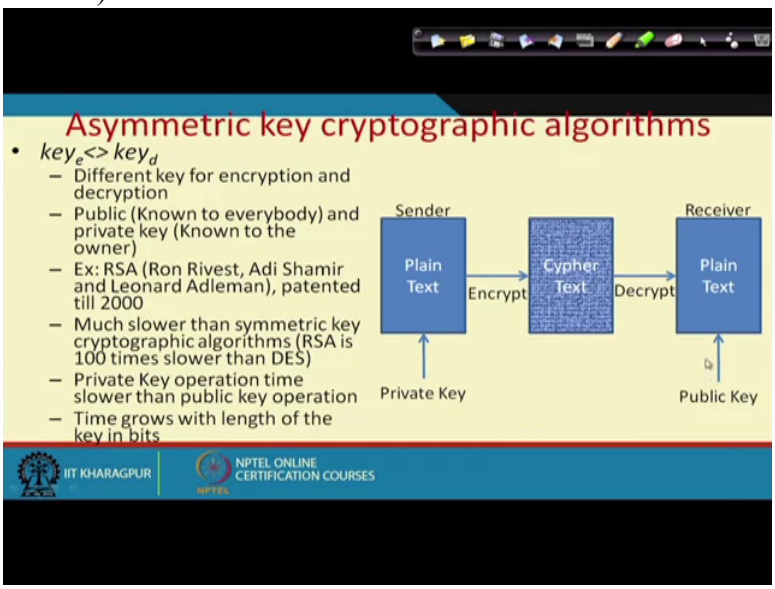
hundred times; this R S A is almost hundred times slower than that of D E S. D E S is a symmetric key algorithm. And there are two operations here, one is private key operation, another is public key operation. Let us first look at this, this private key, public key part operation what we are right now talking about.

First of all sender will be sending the plain text which will be encrypted by its private key. And the keys are generated in the manner so that if something is encrypted with the private key, I mean, if something is encrypted with any of these keys, then using the other one it can be decrypted. So if something is encrypted by, I am sending something to somebody then if my plain text is encoded with my private key which is known to me and I send it over the internet. Whosoever has my public key can actually decrypt it, Ok, or

(Refer Slide Time 11:38)

if somebody has my public key and he encrypts the data with my public key then when it comes back to me, I can only decrypt it using my private key. So which means the problem of key transfer, that we were talking about, while discussing this your symmetric key cryptographic algorithm this,

(Refer Slide Time 12:10)



the text that is coming to me, because everybody knows my public key, so they can actually encrypt it using this public key and send the cipher text back to me and nobody else in the world know my

(Refer Slide Time 12:28)

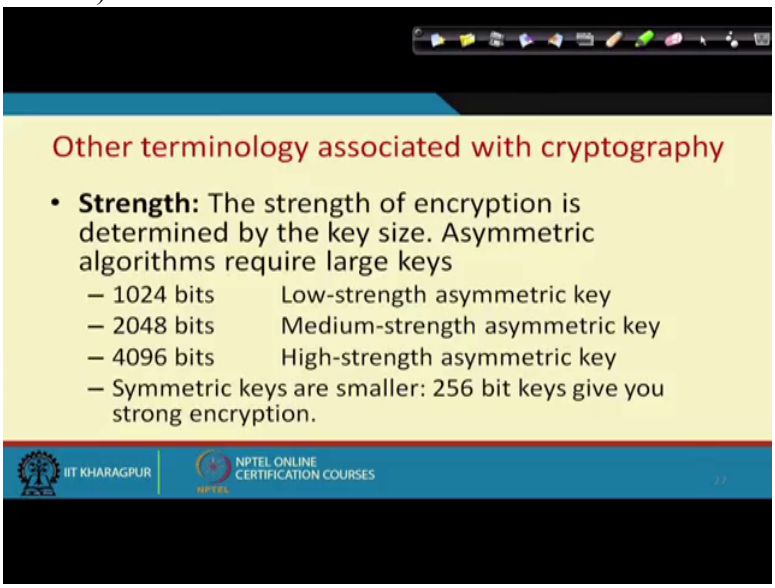
private key. So therefore I am the only person who can decrypt it. So if the key gets generated at the receiver's end, then I will be able to, all over the internet nobody else can know the key if he, if that person, the sender is using my public key to encrypt it. I am the only person who can actually decrypt it and read the text. So this is how my key distribution problem is solved. Because even if the algorithm, this particular algorithm is very slow, because the key will be usually small size, let's say 1 0 2 4 bit etc, so it will not be very large.

So therefore a small chunk of data at least we can afford to send that small chunk of data, which is my key over this. So this is one application of symmetric key. Let us look another application. Suppose I am sending some data using the private key. If I encrypt the data using my private key it is true that everybody in the world can actually decrypt and see it because my public key is publicly available. But the benefit here is because the data is getting decrypted my public key it is sure that I am the person who is sending it. So this non-repudiation issue is also resolved. So I can never say that I have not sent it. Ok, but we are going to know little bit more about them later.

So as I was telling you this particular algorithm, this asymmetric key cryptographic algorithms, they are actually much slower and of these, because once you have a private key operation and once you have a public key operation whether for encryption or decryption, so of these two operations this private key operation time is slower than this public key operation time. And as the key size goes up so also the time.

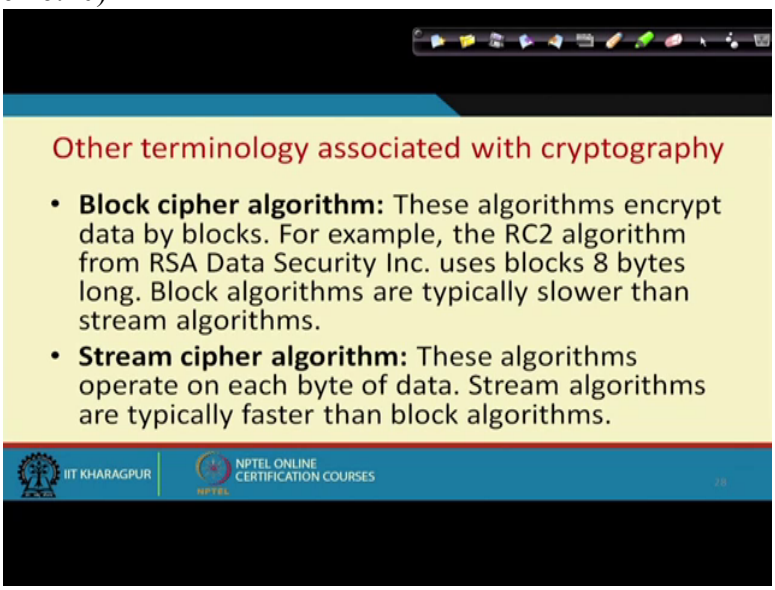
So there are certain other terminologies which come in the context of cryptography. That is what I have already mentioned the key size, that is a very important factor, as the key size grows then the time to do the operation also increase but this increase in operation, the time for encryption also leads to the benefit that it becomes even more difficult to decrypt that. Even if attacker tries, as the key size grows, because he himself has to carry out huge number of operations possibly in years together he will be keep trying various options to decrypt it. And that is how if you are sending a business document over the internet which is secure because the time would be so huge to decrypt it so that by the time he has decrypted, let's say after 2 years, 3 years, 5 years, it is actually useless piece of information. So the strength of the key,

(Refer Slide Time 15:58)



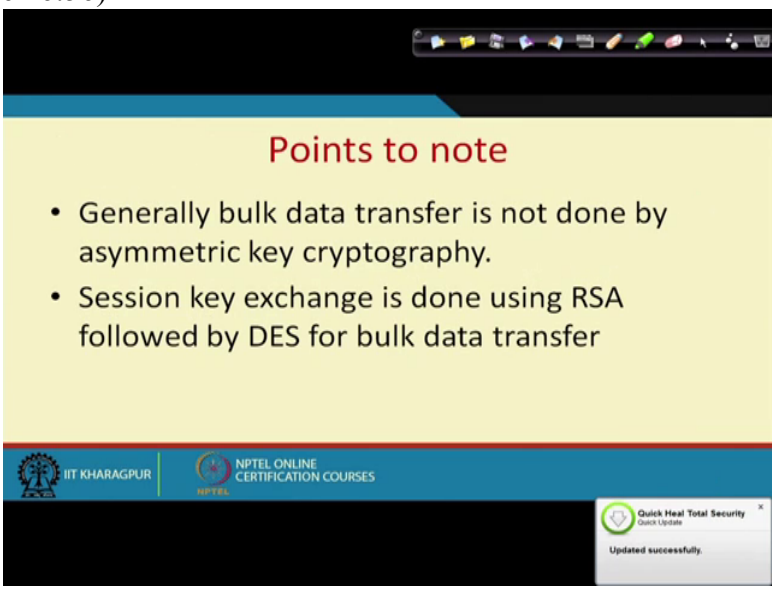
the size of the key basically is associated with the strength of the algorithm. So the strength of encryption is determined by key size. Asymmetric algorithms require large keys and symmetrical key algorithms usually require smaller key sizes to give you strong encryption.

(Refer Slide Time 16:20)



Then the algorithms which belong to the cryptography can be broadly classified into two categories, one is your block cipher algorithm, second one is your stream cipher algorithm. So these block cipher algorithms actually encrypt the data by blocks. Then your stream cipher algorithms actually operate on each byte of data. The stream cipher algorithms are typically faster than the block algorithms.

(Refer Slide Time 16:56)

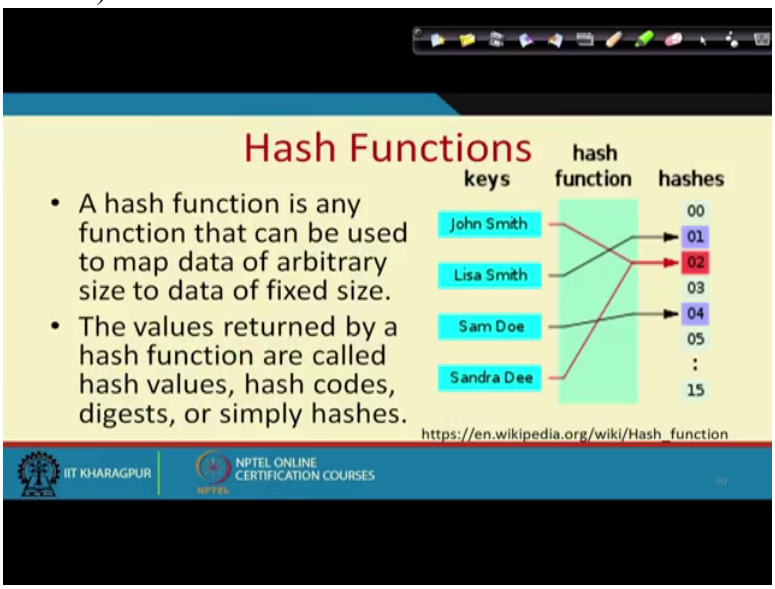


Then this is what I have been telling you so long. This block usually for exchanging the keys, we use that expensive public and private key pairs which is called asymmetric key algorithm but once the

(Refer Slide Time 17:16)

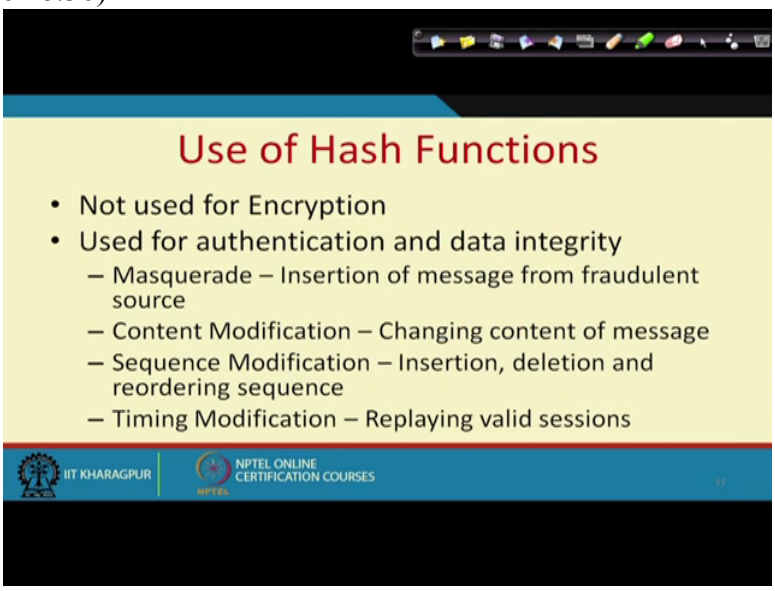
key gets transferred we can use the symmetric key which is now known to both the parties for bulk transfer of the data.

(Refer Slide Time 17:28)



Then another important terminology that we are going to use subsequently is your hash function. This hash function is any function that can be used to map the data of arbitrary size to a fixed size and this value returned by hash function is called hash values, hash codes, digest or simply hashes. For example these different names given here, if they pass through hash function they just produce some hash numbers. And these hash numbers, it is very difficult to find, given this data it is very difficult to find, it is very difficult to reverse from the hashes, from the message digests to get the original data. But they will be actually giving a number, much shorter than the original message.

(Refer Slide Time 18:36)

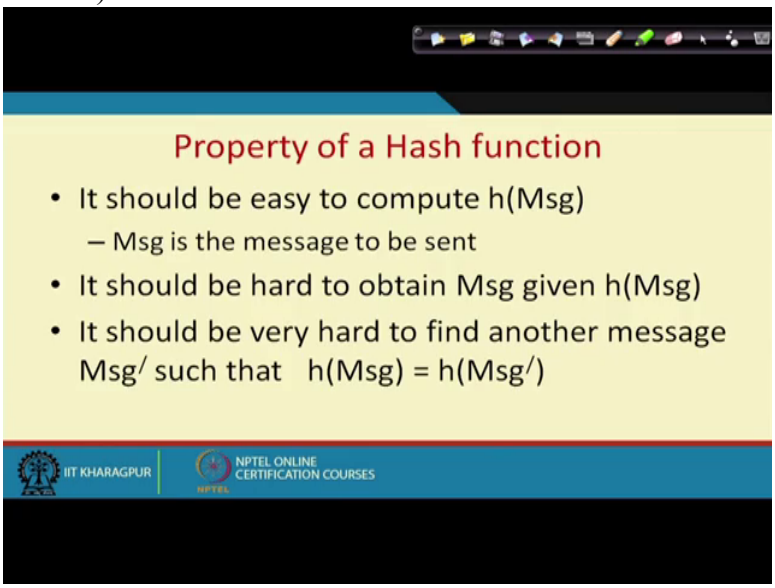


Now what is the use of these hash functions? These hash functions are first of all, we should be sure that, we should remember that hash functions are not for encryption. They are used for authentication and data integrity. And about what is this authentication and data integrity, we have already discussed. So if somebody tries to insert a message from a fraudulent source, if somebody tries to modify the content of the message, of the application data that we are saving, if somebody tries to modify the sequence, what is that sequence? That sequence number is the T C P sequence number. If the sequence numbers are changed then the data, application data will be arranged, at the receiver side it will be arranged in a manner so that it will look irrelevant. Then it also associated with this modifying the time so that, you know, like valid

(Refer Slide Time 19:45)

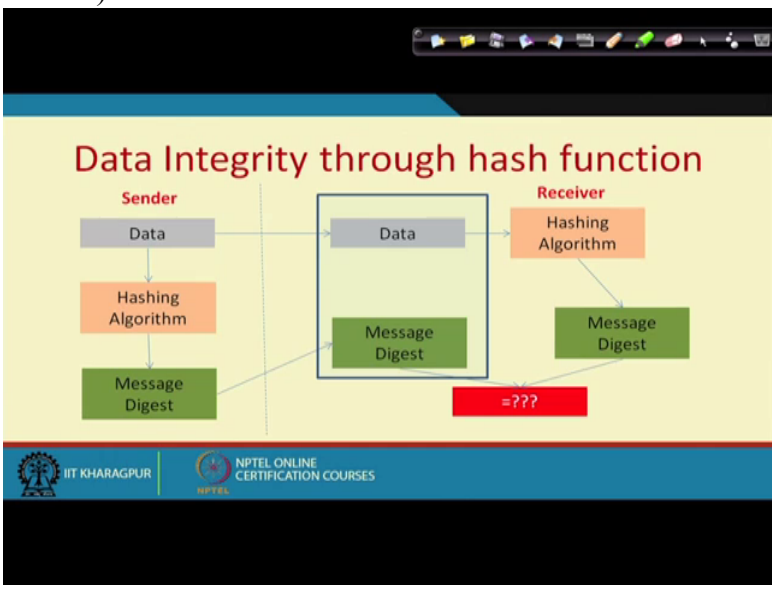
sessions will be, valid session transactions during internet, over the internet can be repeated. Now as I have told you what should be the properties of hash function? It should be easy to compute, it should be hard to obtain the message back given the hash function. And it should be very hard another message so that

(Refer Slide Time 20:14)



both the messages produces the same hash function, it should be very difficult.

(Refer Slide Time 20:25)



So look at this.

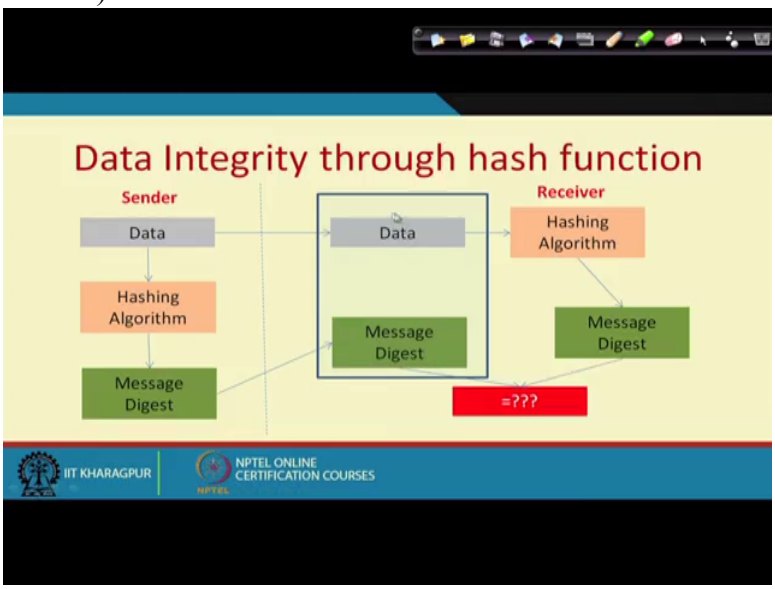
Here we are showing that how exactly data integrity can be maintained using the hash function. Look what the sender will be doing; sender will be using certain hash algorithm to

create the message digest or the hash value. Then what is data integrity? Data integrity is that the data that you are sending is actually the same data which is received by the receiver. The receiver may not receive the same data as we have discussed. The data part can be modified, it can be, the content can be changed, the sequence can be changed and so on. So therefore

(Refer Slide Time 21:14)

both the data as well as the message digest is sent to the receiver. So the receiver

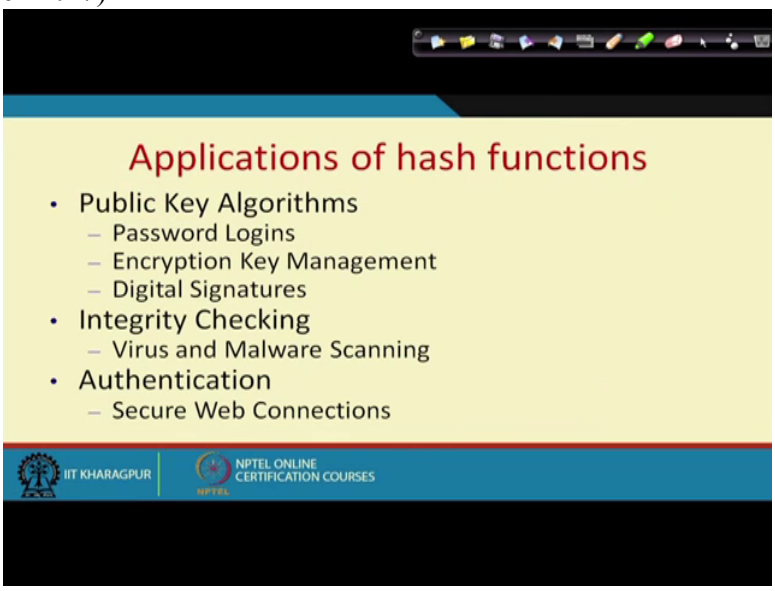
(Refer Slide Time 21:22)



at this end uses this data and the same hashing algorithm to find out the message digest. And this message digest is otherwise sent along with the data. So now at the receiver's end, receiver compares both the data and because it is very hard to get the same hash value for different application data, at least now the receiver is sure it has got because it has got the

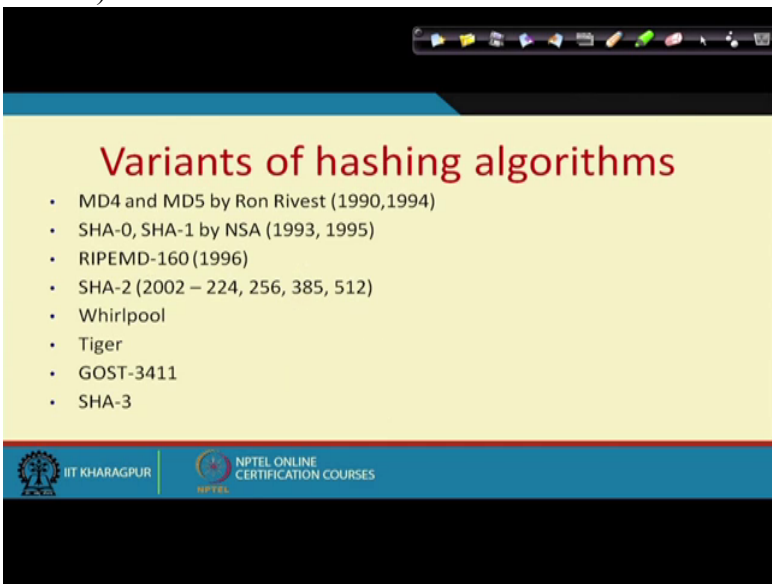
same message digest, now it can be sure that the data that he has received is actually the data which is sent by the sender. On this way, it is not modified.

(Refer Slide Time 22:17)



Then there are various applications of the hash function and in the next lecture possibly we are going to look at some of the applications in the area of public key algorithm and for public key algorithm and later classes we will be looking at its, how it is used for authentication purpose during web transaction.

(Refer Slide Time 22:44)



So besides the, Ok there are various algorithms, hashing algorithms just like your symmetric key algorithm, there are many symmetric key algorithms, also for hashing there are many

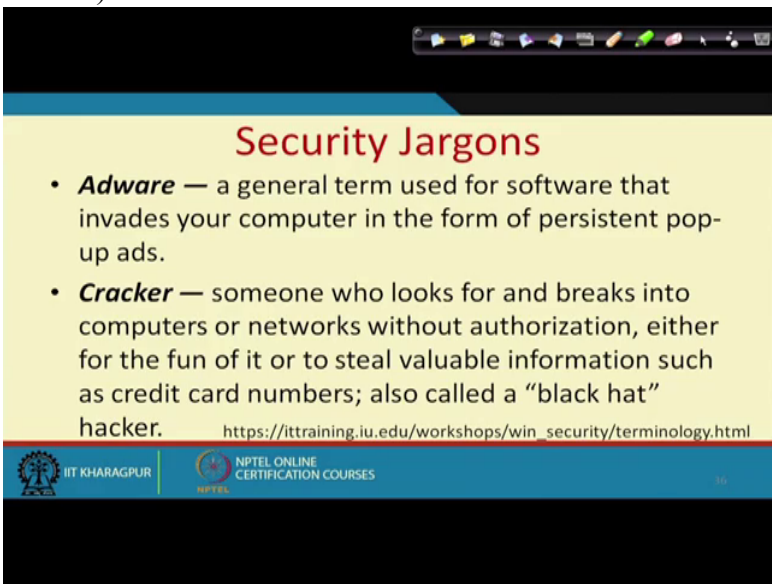
algorithms. They have their respective advantages and disadvantages but we are right now not

(Refer Slide Time 23:02)

going to discuss about them.

In the security

(Refer Slide Time 23:08)



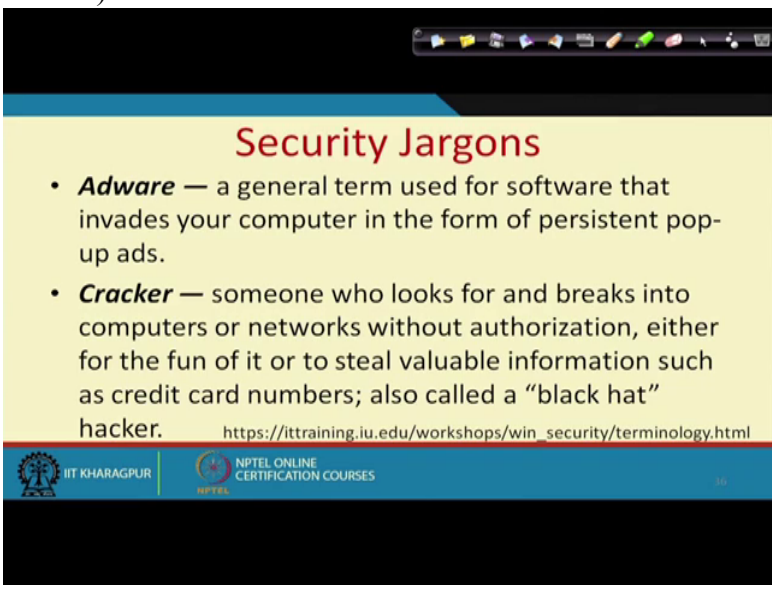
area, there are many jargons which are used by people and

(Refer Slide Time 23:16)

which have almost become standard and referring various types of attacks, people use those terms.

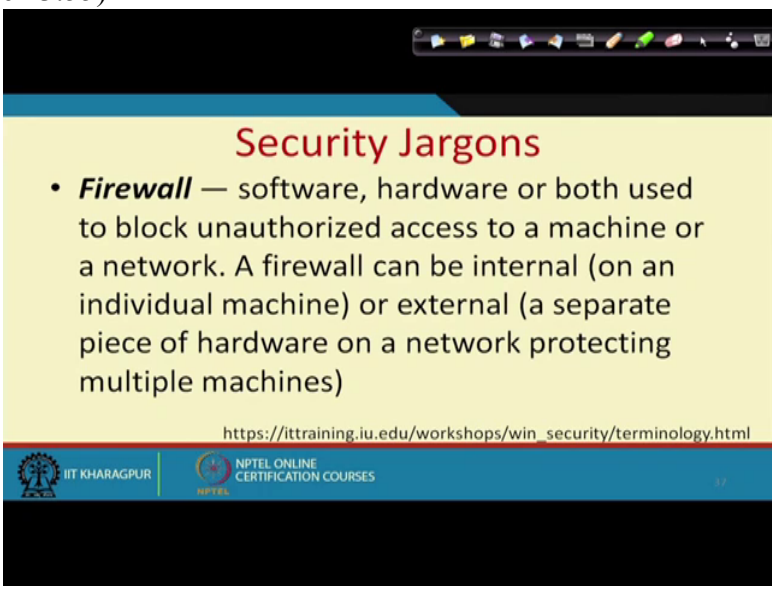
First is your adware. It is a generic

(Refer Slide Time 23:27)



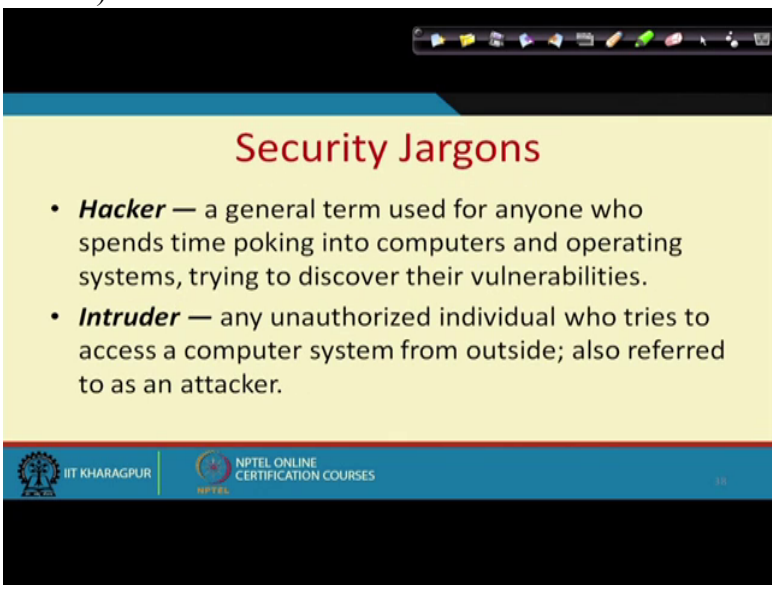
term used for software that invades your computer in the form of persistent popup ads. Then you have a term called cracker. It is someone who looks for and breaks into the computer network without authorization either for fun, or to steal the valuable information such as credit card number etc.

(Refer Slide Time 23:55)



Then very frequently used term is a firewall. It can be either a software or a hardware or can be a combination of both. It can be, it prevents various types of attacks. It can be internal. It can reside on one individual machine or it can be external. It is for the entire network for protecting it, protecting multiple machines at a time. It can be set up in your proxy which means it can be set up in your proxy.

(Refer Slide Time 24:29)



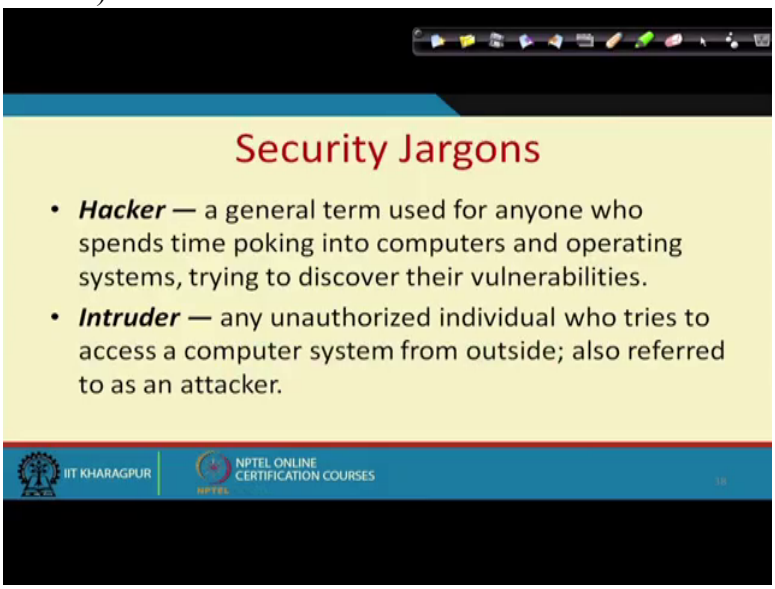
Then hacker is somebody, somebody who spends time poking into the computers and operating systems and tries to discover their vulnerability. So while talking about these hackers

(Refer Slide Time 24:49)

who actually enter into the system and tries to find out its vulnerability, another term comes into mind which is ethical hacking. By ethical, ethical hackers are those hackers who are hired by the companies to do hacking on their own system to find out the security holes. Once these ethical hackers find out the security holes, then the company can take adequate measures to get rid of those security holes.

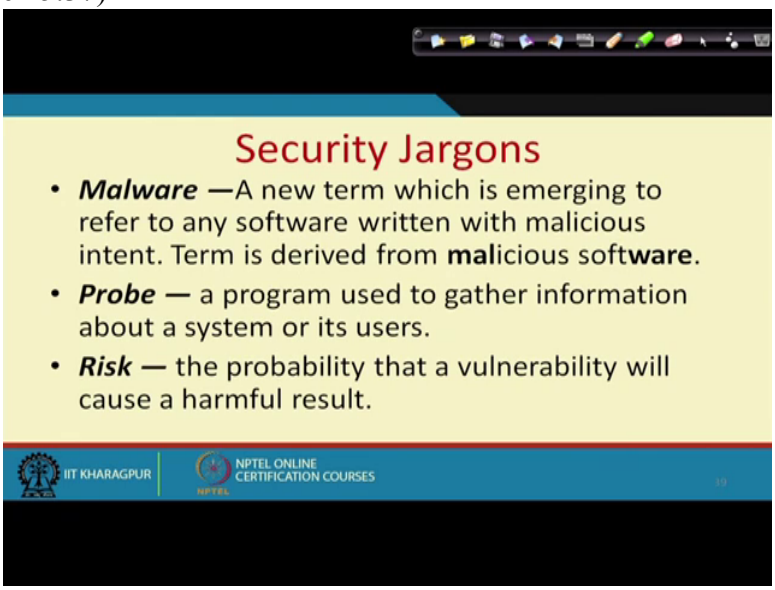
Then the next term is intruder.

(Refer Slide Time 25:28)



This is any unauthorized individual who tries to access a computer system from outside. They are also known as attacker.

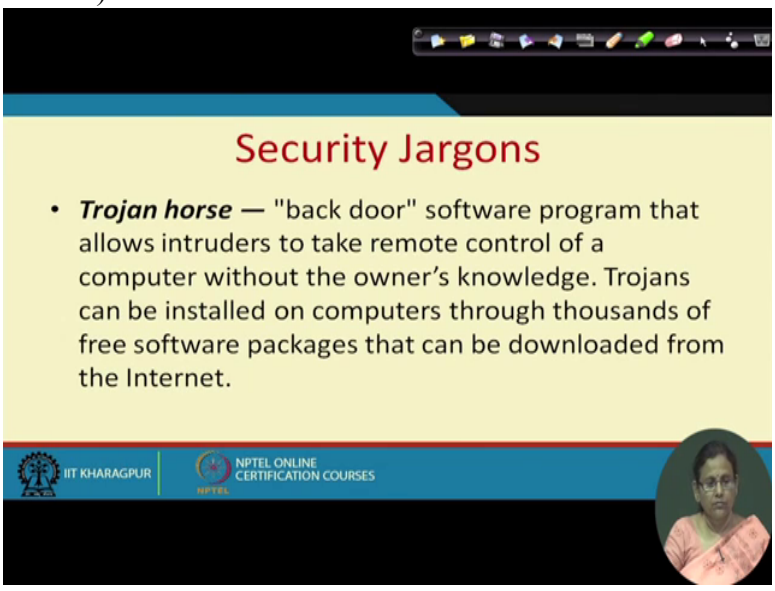
(Refer Slide Time 25:37)



Then next security jargon is your malware. It's the term which is emerging to refer to any software written with malicious intent. This term is derived from using these two terms, malicious and software, they are called malware. Then probe, probe is a program used to gather information about a system or its users.

Next term is your risk. This is the probability that, that vulnerability will actually cause some harmful effect.

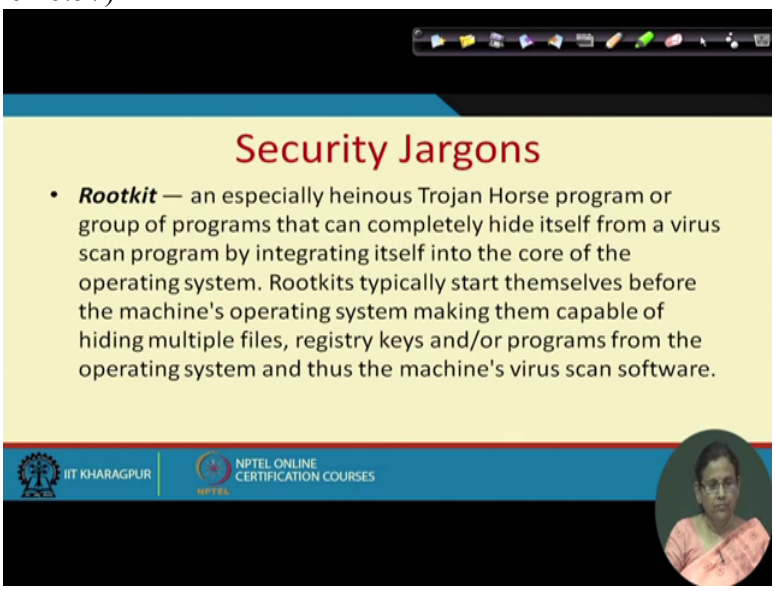
(Refer Slide Time 26:21)



Then next term is your Trojan horse. They are backdoor software programs that allows intruders take remote control of a computer without owner's knowledge. These Trojans can be installed on computers through thousands of free software packages that can be downloaded

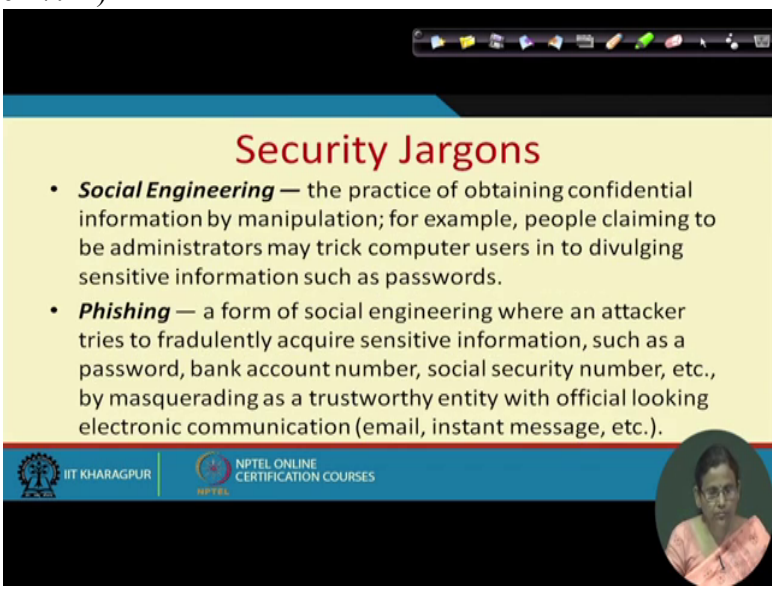
from the internet so the computer system has to be prepared so that it does not allow the outside software packages to get installed into itself.

(Refer Slide Time 26:57)



Then rootkit, it is again, it is an especially heinous Trojan horse program or a group of programs that can be, that can completely hide itself from virus scan program by integrating itself into core of the operating system. This rootkit typically starts themselves before the machine operating system is capable of handling files and these such programs are capable of hiding multiple files, registry keys or programs from the operating system thus the machine's virus can, software can also be hidden.

(Refer Slide Time 27:44)

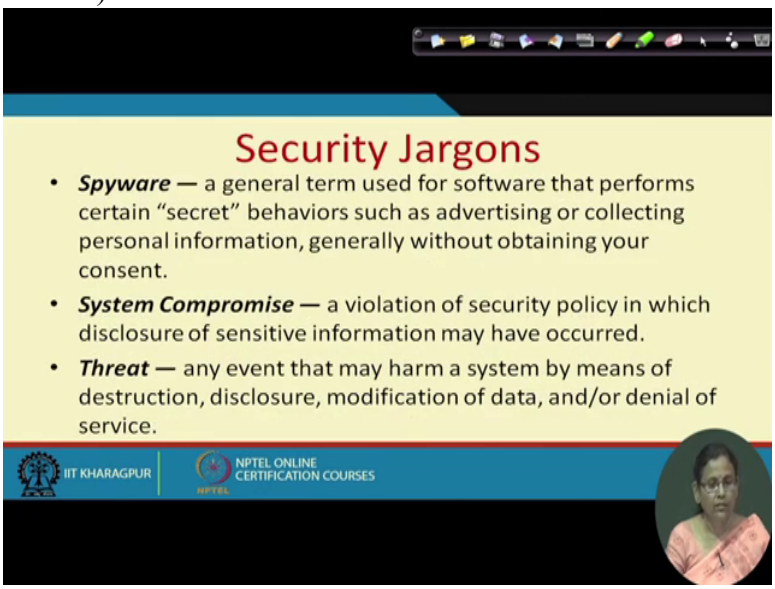


Then another associated term is called social engineering. So it is the practice of obtaining confidential information by manipulation. For example the people claiming to be

administrator may trick the computer user into divulging sensitive information such as passwords.

Then another related term is called phishing. It is a form of social engineering where an attacker tries to fraudulently acquire sensitive information such as password, bank account number, social security number etc. behaving as trustworthy official entity.

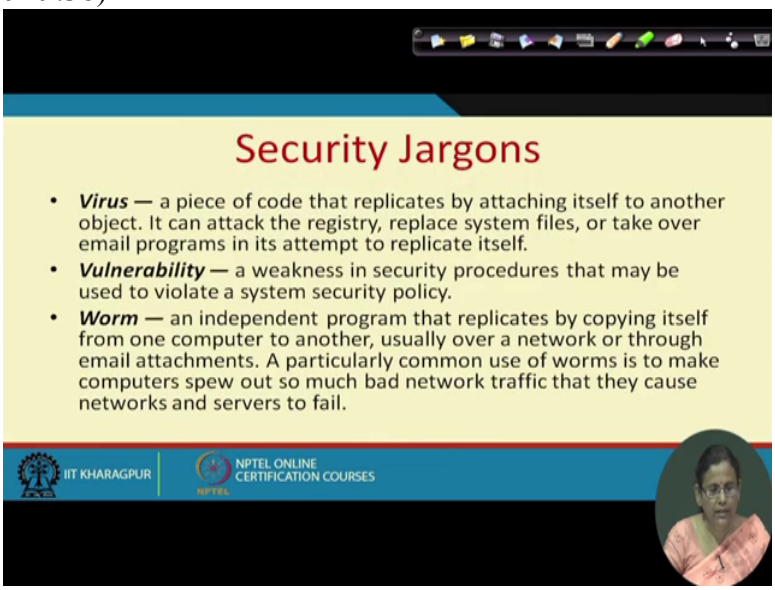
(Refer Slide Time 28:28)



Then next is your spyware. This is the term used for software that performs certain secret behavior such as advertising, collection of personal information without obtaining your consent. So as I have told you while we were talking about this marketing data collection that companies do collect the data about the user's behavior. However collecting such data without user's permission is not legal. So the software which illegally does this is called spyware.

Then another term, associated term is your system compromise. It is about the violation of the security policy in which the disclosure of sensitive information may have occurred. Then next one, term is a threat. Any event that may harm the system by means of destruction, disclosure or modification of data and/or denial of service is called a, called some threat.

(Refer Slide Time 29:38)



Next virus is something which all of us know. It is a piece of code that replicates by attacking itself to another object. It can attack the registry place of system files and your email program and so on. And it has habit of replicating itself. It goes from one person to the other; that is how it is called virus. Then vulnerability I think we have already discussed, it is the weakness of the security procedure that may be used to violate the system security policy. Then next is a worm. It is an independent program that replicates by copying itself from one computer to another usually over a network or through email attachments. A particular common use of worm is to make computers spew out so much that bad network traffic that they cause network and server to become very slow, sometimes it, the network actually

(Refer Slide Time 30:35)

becomes congested. So with this we finish this lecture and next class onwards we are going to talk about more on security stuffs, thank you very much.