CS 240: Data Encryption Transcript

[00:00:00] **INSTRUCTOR:** The next core security concept that we want to talk about is data encryption. Data encryption is all about keeping your data private and storing it in such a way that people that shouldn't access the data can't read it.

[00:00:17] So if we want to protect our data from others, we encrypt it. Encryption is the process of encoding the data so that only people that are authorized can read it. And the opposite of encryption is decryption so that when we decrypt data that has been encrypted, when we decrypt it, that's the process of decoding it so that we get back the original data in its original form.

Start visual description. Slide titled Encryption/Decryption. Text reads:

- *In Cryptography, Encryption is the process of encoding data so that only authorized parties can read it. Decryption is the process of decoding data back to its original form.*
- *Plaintext => the data you are trying to protect*
    - *"Hello, Alice!"*
- *Ciphertext => the data in its encrypted form. "Cipher" is another name for an encryption algorithm*
    - *"6$8xF9q37^&?"*
- *Key => a piece of information (sequence of bits) used as input to a cryptographic algorithm to encode or decode data (often stored in a file)*
- *Key Size => The number of bits in the key*
    - *E.g., 128 bits, 192 bits, 256 bits, 1024 bits, 2048 bits*
    - *The bigger the key size, the more difficult it will be to crack the encryption*

End visual description.

[00:00:43] And so in security, when we talk about encryption, there's a couple of terms that are important. One is plain text. We've talked about that already. So plain text is simply unencrypted data. For example, we said that we should not store passwords in a database in plain text form.

[00:00:59] And so we decided that we should hash them. So once they've been hashed, they're no longer in plain text. The next term is ciphertext. Now the word cipher is really just another name for an encryption algorithm.

[00:01:12] And so ciphertext is really just data that has been encrypted. Now we use the word text. It's not really text, right? We're still dealing with just bytes of data, but for some reason they use the word text in there.

[00:01:25] So unencrypted data and then encrypted data is what we're talking about. So here you can see hello Alice is the plain text and then the encrypted form of that is a random looking string. So anybody that got their hands on that encrypted string wouldn't have any clue what it says.

[00:01:43] Another term we use a lot when we talk about encryption algorithms is the word key. So usually when you run an encryption algorithm, you have to pass in a key, which is a secret piece of information that is needed to actually encrypt or decrypt the data.

[00:02:01] So think of the key as just being an input to the encryption algorithm. So the inputs to the encryption algorithm are the data that's being encrypted and then the key. And you can kind of think of a key as like a lock on a door, right?

[00:02:17] You have to have a key to open the lock on a door. You're not going to get in unless you have the key. So by the same token you can't encrypt or decrypt data unless you have a key. When we talk about a key, it's just a sequence of bits.

[00:02:28] So it's yet another array of bytes and it's one of the inputs to our algorithm. We also talk a lot about key sizes when we talk about encryption. So the key size would just be the number of bits that make up the key.

[00:02:44] So how big is the key, right? How many bits is it? So it could be 128 bits, 192, 256, or much, much bigger. And in a lot of applications we use keys that are like 2048 bits or 4096. And so key size is important, because typically the bigger your key is, the harder it's going to be to hack into your data.

[00:03:06] It's going to be harder to decrypt the data. And so big keys are more secure than smaller keys, is the idea. So in order to decrypt and decrypt data, we need an encryption algorithm. And there are a lot of these out there, some of which are far more commonly used than others.

[00:03:30] And so we'll talk about what the most commonly used ones are. But there's two different types of encryption algorithms. So we talk about symmetric key algorithms. And we also talk about asymmetric key algorithms.

Start visual description. Slide titled Encryption Algorithms. Text reads:

- *Modern encryption algorithms are divided into two categories*
  - *"Symmetric Key" (or "Secret Key") algorithms*
  - *"Asymmetric Key" (or "Public Key") algorithms*

End visual description.

[00:03:44] And both are very important. Symmetric key algorithms are algorithms such that the key that you pass into encrypt some data is the same key that you would use to decrypt the data. So the same key is used to encrypt and decrypt the data.

[00:04:00] An asymmetric key algorithm is an algorithm where there's two different keys. One of the keys is used to encrypt the data. And the other key is used to decrypt it. And because we're using different keys for encryption and decryption, then they call that asymmetric key.

[00:04:15] Now there's other words for these. Sometimes symmetric key algorithms are called secret key and asymmetric key algorithms are called public key algorithms.