CSCD27 – Computer and Network Security

1 Main Notes

Contents

2 Cryptography

What is Cryptography

Cryptography is the subject that introduces tools that we use to secure communications. Often times the channel we use to send messages is insecure, whether over a road or over the Internet. We use cryptography to secure these channels primarily by obfuscating data to ensure confidentiality (CIA). Because we are just obfuscating the message, cryptography cannot address integrity or availability.

Kinds of Cryptography Work

There are two kinds of cryptography fields, one is the pure mathematical side building the individual tools. Then there is a computer science/technical side that utilizes these tools to build crypto protocols.

2.1 Classical Cryptography

Plaintext

Original message intended for author and recipient. This is the message that we want to secure.

Ciphertext

Encrypted message.

Encryption

The process of turning plaintext into ciphertext.

Decryption

The process of turning cipertext into plaintext.

Cryptographic algorithm

The algorithm that is used to do encryption and decryption.

Cryptographic key

Input used by algorithm to perform the actions of encryption and decryption.

Shift Cipher

A **shift cipher** is one where you shift the characters in the alphabet by a fixed offset. Mathematically, shift ciphers can be expressed using $a + b \mod 26$ since if our shift offset goes past 26, we wrap back around.

The shift cipher is the oldest kind of cipher known to have been created. It was first implemented in the Caesar Cipher. With the Caesar cipher, the key is the shift offset. That is all fine and dandy, but we know that this is obviously trivial to crack. But how can we measure this? This is where we introduce key entropy.

N-Bit security entropy

A measure in which we measure the strength of the algorithm. It is the number of bits to encode the number of possible keys.

So for example in the Caesar cipher, with 25 possible rotations (not counting a full rotation which is dumb) you need 5 bits. This doesn't seem like a lot, and it isn't. While we will see that entropy is not exactly the be-all-end-all of measuring how good a cipher is, it is a useful benchmark as it directly impacts how easily a cipher is brute forced.

Quick Facts – Caesar Cipher

Type: Monoalphabetic Shift Cipher
Key: Shift offset
N-bit security entropy: 5 bits

What is considered good entropy?

100 bits is the minimum threshold to not be considered a weak key.

In the example big picture, an integrity attack can be where Mallory takes the message and sends it to someone else other than Bob, or resends it to Bob later.

Kerckhoffs' Principle

Kerckhoffs' Principle can be summed up as: "The enemy knows the system", or in other words, there is no security by obscurity. A cryptography algorithm should not rely on the fact the algorithm itself is secret. This only doesn't apply to the key.

Types of ciphersystem attacks

Listed are some of the main ways that ciphersystems are attacked. Note that they can be mixed together to form a more effective attack.
Brute force: for n keys it takes on average $n/2$ attempts to crack.
Ciphertext Known: (you are given) One or more random ciphertexts are known. Generally this attack occurs using something like statistical analysis.
Known Plaintext: (you are given) You know one or several plaintext(s) with it's associated ciphertext.
Chosen Plaintext: (you get to interact with the system) You feed the cipher a plaintext and get the associated ciphertext.
Chosen Ciphertext: (you get to interact with the system) You feed the cipher a cipher text and get the associated plaintext.

So how can we apply these attacks to the Caesar Cipher? Brute forcing is very easy, you don't even need a computer. You just need to recognize what cipher is at play. Ciphertext Known for Caeser Cipher, you can try can try to reverse engineer the cipher by analysis, as opposed to brute forcing. So if you know something about the nature of the plaintext, you can do this.
Chosen Plaintext and Ciphertext are really just mirrors. So for Caesar cipher for chosen plaintext, you input a plaintext character and count the offset. For Chosen Ciphertext you do the same but with the ciphertext character

(so basically in this case it looks identical).

So obviously it isn't great. Aside from the weak key permitting brute force, one can also analyze the structural weakness of the cipher. For example, with 'e' being one of the most common letters in the alphabet, you can easily sniff out or guess which character is meant to substitute 'e'.

Monoalphabetic Cipher

A cipher that does not change the frequency of letters in a message. This is generally a bad idea. They are vulnerable to statistical analysis. The Caesar Cipher is one of them.

Substitution is another example of a monoalphabetic ciper (along with Caesar).

Quick Facts – Substitution Cipher

Type: Monoalphabetic Cipher
Key: The permutation mapping
N-bit security entropy: 26! bits

Substitution cipher is one where you create a random permutation of the alphabet and use that. So if our alphabet was (a, b, c), then we could have (b, c, a) and we would subtitute b for a, c for b, etc.
As we can see by the much larger entropy, it is far more resilient against brute force. However, as hinted earlier, higher entropy does not necessarily guarantee more security. A substitution cipher is vulnerable still to the other four attacks. In fact, you can combine known plaintext with bruteforce, it is common to mix them.

So the opposite of a monoalphabetic cipher is a polyalphabetic cipher.

Polyalphabetic Cipher

A polyalphabetic cipher is one where we do change the fequency of letters in a message. This makes the cipher more resilient against analysis based attacks.

One example for a polyalphabetic cipher is a Vigenere cipher. Instead of a fixed permutation, we have a phrase or word as our key instead of a number or set of permutations. So for example, if our word is BUTTERDOG, the first letter would be offset by 2, the second would be offset by 20, so on and so forth. Once we have finished typing out BUTTERDOG, we restart again until we have exhausted the plaintext.

Quick Facts – Vigenere/Renaissance Cipher

Type: Polyalphabetic Shift Cipher
Key: A word (or a phrase with no spaces or whatever you want)
N-bit security entropy: 26^n

A big enough n makes brute force impossible for Vigenere. Statistical analysis is also dependent on length and the quality of the key. In an attack what you could do is try to reduce the search area to make brute force more viable. The other attacks are still possible if not easy.
So as we can see, if the key has an insufficient length, or has poor form, then Vigenere is unlikely to be very strong. But then what if we solved both of those in one shot, or in other words, in One Time Pad.

Quick Facts – Quick Pad

Type: Polyalphabetic Shift Cipher
Key: An infinite string containing randomly generated characters
N-bit security entropy: ∞

OTP is the best cipher eva. The reason why its called One Time Pad, you can only use the key once. This helps build resilience against attacks in where an attacker has access to the cipher, or has access to ciphertexts and plaintexts. In this case we use XOR the two strings together to compute the ciphertext, rather than using the key's characters as a shift offset. Using the random string, we eliminate statistical analysis. And since we use the key only once, known plaintext is no good. Problem with this is that we have to constantly exchange keys. For one our key has an infinite size. Second, our key is only valid for a single exchange, so every message sent must get a new key. However, there are ways to improve OTP and make them more viable for real world applications. Many ciphersystems base themselves on OTP.

One suggestion given during lecture was to use the previous message as the encryption key. It is one time, mostly random, and we have to send a message either way. Only issue is that the chain has to be seeded from somewhere and requires that initial exchange of keys.

Random single-use keys with high entropy are best.

Randomness makes statistical analysis extremely difficult/impossible, higher entropy keys make brute forcing unviable. The key being single-use makes known plaintext-ciphertext attacks impossible.

Transposition Cipher is where we have a message of a fixed size, and each message we shuffle around characters in the messages. The key is the set of permutations. Brute force again depends on the entropy. You can still do some analysis if you have a known ciphertext. Known Plaintext you can reduce size, Chosen you send the alphabet and match the letters.

Quick Facts – Transposition Cipher

Type: Polyalphabetic Cipher
Key: A permutation of the order of the characters
N-Bit security entropy: $n!$ where n is the length of the fixed length message.

Pillars of Modern Cryptography

The pillars of modern cryptography are: Diffusion, Confusion and Randomization. Diffusion is like the game of guessing which cup has the ball, you want to mix up the characters in the string and make sure that changes are not local. An example is transposition. Confusion is replacing a symbol with another but keeping this relationship as opaque as possible. Randomization is not quite what you think it is, it is the feature where the same text encrypted multiple times does not generate the same ciphertext. Confusion and Randomization are good enough by themselves and are used to build a stream cipher. Diffusion and Confusion can be mixed together to build a block cipher.

3 Core Concepts

Definition - CIA

An acronym for the three pillars of security, confidentiality, integrity and availability. For example, confidentiality means only the student can see their own grade, integrity is only the professor can enter a grade, availability just refers to keeping important systems online like 2FA. There are sub properties, see slides for Venn diagram.

Something curious about the pillars of security is that sometimes some of them are undesirable. For example something like confidentiality and availability might be targets. Additionally, they can even conflict with each other. For example we have voting. For confidentiality, you want your vote to be secret so that you can express yourself freely. For integrity, you want to keep someone from voting twice, or having their vote changed, or pretending their vote never happened. And for availability, you want everyone who is able to vote to do so. Integrity and confidentiality in this case begin to clash a bit.

Anonymity vs Pseudonymity

Pseudonymity is the state of something like Discord or Reddit. You have a username (likely not your real name) that isn't directly named after you. However, as you show yourself through these platforms, people can begin to triangulate who this account likely belongs to, thus potentially revealing yourself.
True anonymity means that it is impossible to trace something back to your person through any means.

4 Glossary

Safety – Concept of a program where reasonable input for reasonable output, or correctness, or the program does what is intended.
Security – Concept of a program where unreasonable inputs get reasonable outputs.
Threat – Possibility of damage, can be tangible, possible, whatever. A threat can be accepted or not accepted. Rejected threats are ones that are not cost-effective or low possibility.
Vulnerability – A weakness in the system, but not necessarily something that can be exploited or no exploit has been discovered yet.

A vulnerability enables a threat, and an attack/exploit takes advantage of a vulnerability. There is a clear differentiation between an attack and a vulnerability. An attack realizes a threat.
Countermeasure – something done to disable an attack, limit the possibility or consequences of damage. So this can be like a patch or workaround. It mitigates a thread or removes/reduces a vulnerability
Interception – A kind of attack where an attacker obtains and reads the message. (Confidentiality)
Modification – A kind of attack where an attacker changes the message (Integrity)
Fabrication – A kind of attack where an attacker injects a message (Integrity)
Interruption – A kind of attack where an attacker disrupts communications (Availability)