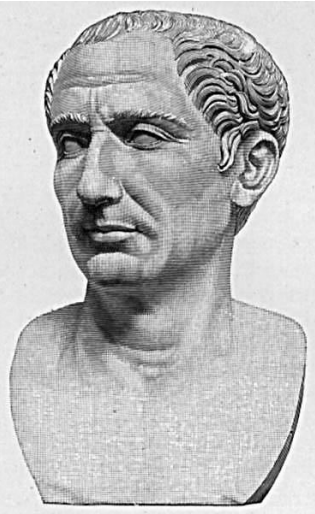
Beveiliging – Les 1

Cryptografie is de studie van het verbergen of versleutelen van gegevens. Dit is nodig als we gegevens willen doorsturen of gevoelige gegevens (zoals paswoorden of kredietkaardnummers) willen opslaan. Gegevens die worden doorgestuurd zijn eenvoudig af te luisteren en worden best steeds versleuteld doorgestuurd. Het is dus belangrijk om gegevens onleesbaar te maken voor ze worden verzonden (= versleutelen of *to encrypt*) en terug leesbaar te maken (= ontsleutelen of *to decrypt*) bij aankomst.

In deze les bekijken we enkele cryptografische technieken en principes als inleiding. De technieken in deze les zijn heel eenvoudig en zijn niet sterk genoeg om gegevens mee te versleutelen, maar illustreren wel enkele cryptografische principes. De meest gebruikte methodes (AES en RSA) worden in volgende lessen besproken.

Caesar cipher (Caesar versleuteling)



Om een tekst te versleutelen (encryptying) zouden we elk karakter in de tekst een plaats in het aflabet kunnen opschuiven. Je hebt deze techniek misschien wel eens gebruikt als kind. Als we deze tekst zouden willen versleutelen:

```
Hallo Bob
```

Kunnen we de karakter bijvoorbeeld 7 plaatsen opschuiven in het alfabet. H wordt dan bijvoorbeeld O. De a wordt dan h. De versleutelde tekst, of **cijfertekst**, is dan:

```
Ohssv Ivi
```

We verplaatsen de karakters zo (bovenaan is de input, onderaan de output):

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G

Om de tekst terug te ontcijferen (decrypting) moeten we de karakters gewoon terug naar de andere kant opschuiven, ofwel in de bovenstaande aflabetten de onderste regel als input gebruiken.

Deze versleuteling is niet erg veilig. Als een klein beetje slimme aanvaller (attacker) de tekst zal analyseren zal hij merken dat sommige letters meer voorkomen dan andere. Hieruit kan hij afleiden dat de tekst niet sterk versleuteld is. De meest voorkomende letter in het Nederlands is

de “e” (gemiddeld 1 op 5, zie https://onzetaal.nl/taaladvies/letterfrequentie-in-het-nederlands). Het meest voorkomende letter in een tekst versleuteld via het Caesar cipher zal waarschijnlijk een versleutelde “e” zijn en van daaruit kan dan de verschuiving worden berekend. Uiteraard werkt het gemakkelijker met een langere tekst (in ons voorbeeld komt er toevallig geen “e” in de tekst voor). Deze methode maakt gebruik van letter-frequentie-analyse. Een andere methode is woord-frequentie-analyse: analyse door de woorden te rangschikken volgens voorkomen in een gemiddelde tekst. Sommige woorden komen meer voor dan andere, denk aan de woorden “de”, “het”, “een”, “van”, “dat”, “en”, …

Een andere manier om een tekst te ontcijferen is om gewoon de 25 verschuivingen te gaan berekenen en kijken welke tekst leesbaar blijkt te zijn. Dit wordt een *brute force attack* genoemd: al de mogelijkheden worden geprobeerd.

Oefening: Caesar cipher

a. Versleutelen

Maak een programma waarin je een tekst kan ingeven en een getal (de verschuiving). De tekst wordt vervolgens versleuteld met het Caesar cipher.

b. Ontcijferen met brute force

Maak een programma waarin je een versleutelde boodschap met het Caesar cipher kan zetten, het programma berekent al de 25 mogelijke verschuivingen.

Extra: ontcijferen met eenvoudige letter-frequentie-analyse.

Kan je het programma een gok laten maken wat de meest waarschijnlijke verschuiving is door de meest voorkomende letter te berekenen?

Mono-alphabetic cipher

In plaats van het alfabet te verschuiven kunnen we elke letter toewijzen aan een willkeurige andere letter. We krijgen zo een sterkere versleuteling die moeilijker met brute force te kraken is, maar nog wel gemakkelijk met letter-frequentie-analyse.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
C	Y	R	J	A	H	X	L	K	S	Z	I	P	Q	U	M	F	B	N	O	G	W	T	E	V	D

Mono-alphabetic cipher ontsleutelen

Als je de sleutel hebt om tekst te versleutelen heb je ook de sleutel op de tekst te ontsleutelen. Elke letter in de onderste rij staat voor een verschuiving, bijvoorbeeld A -> C. Om te ontsleutelen moet elke C terug 2 plaatsen naar links worden opgeschoven en draaien we de sleutel dus om: C -> A.

De versleutel-sleutel en ontsleutel-sleutel zijn dus hetzelfde of kunnen van elkaar worden afgeleid. Een versleutelingstechniek waarbij de sleutel om te versleutelen en ontsleutelen dezelfde is (of van elkaar kunnen worden afgeleid) noemt men een **symmetrische versleuteling**.

Het kan ook zijn dat de sleutel om te ontsleutelen niet van de encryption key kan worden afgeleid, dan spreken we van een **assymmetrische versleuteling**. Hierover later meer.

Oefening: Mono-alphabetic cipher

Maak een programma waarmee je een sleutel kan genereren en bewaren (in een object). Voorzie een optie om te versleutelen en te ontsleutelen.