Julius Caesar was de eerste die een geheimschrift maakte door met een eenvoudige formule de letters van het alfabet door elkaar te schudden. Door de rekenkracht van de moderne computers zijn zulke eenvoudige geheimschriften erg onveilig geworden.

Eenvoudige geheimschriften

Willem van Ravenstein

Om veilig boodschappen naar zijn legers te versturen gebruikte de Romeinse keizer Julius Caesar een geheimschrift. Met zijn generaals sprak hij af dat hij alle letters in het alfabet drie plaatsen naar rechts zou opschuiven. De A werd dus een D, de B een E, de C een F, enzovoort. De boodschap 'KOM NAAR ROME' werd op deze manier 'MRP QDDU URPH.' Je zult begrijpen dat de vijanden van Julius Caesar niet erg veel wijzer werden als ze zo'n boodschap onderschepten.

Als we de letters in het alfabet allemaal een nummer geven (A=0, B=1, C=2, enzovoort), dan kunnen we dit geheimschrift ook voorstellen met de formule $F(x) = x + 3$. Hier staat dat je alle letters in het alfabet drie plaatsen moet opschuiven. Dit klopt niet helemaal, want X = 23 schuift op naar A = 0 en Y = 24 naar B = 1. Dit betekent dat we *modulo* 26 moeten rekenen: we rekenen alleen met de getallen 0 tot en met 25, en na de 25 komt in plaats van 26 de 0 (zie ook pagina 10 - 13 van het decembernummer).

In de 'Kleine Nootjes' op pagina 2 en 3 kom je een geheimschrift tegen waarin de A vervangen wordt door een Z, de B door een Y, de C door een X, enzovoort. Dit is het zogenaamde Atbash-geheimschrift, dat van joodse oorsprong is. Bij dit geheimschrift hoort ook een formule, namelijk $F(x) = 25 - x$.

Vercijferen met een formule

In de twee genoemde geheimschriften worden de letters van het alfabet met behulp van een formule door elkaar gehusseld. Een algemenere manier om dit te doen is met de formule:

$$F(x) = ax + b \pmod{26}$$

Hier is x het rangnummer van een letter (A=0, B=1, C=2, enzovoort) en zijn a en b getallen die je zelf mag kiezen. Kies je $a = 1$ en $b = 3$, dan krijg je het geheimschrift van Julius Caesar. Het Atbash-geheimschrift krijg je met $a = -1$ en $b = 25$. Kies je $a = 11$ en $b = 17$, dan krijg je weer een ander geheimschrift. Je kunt dan gewoon uitrekenen waar elke letter op afgebeeld wordt. De letter P = 15 wordt bijvoorbeeld afgebeeld op de letter met rangnummer $11 \times 15 + 17 = 182$ en modulo 26 is dit 0 = A. Het zinnetje: "Pythagoras is een tijdschrift voor jongeren" wordt vercijferd als: AVSQRFPWRH BH JJE SBMYHNQWBUS

OPPW MPEFJWJE. Heb je Internet, dan kun je op www.crypto.club.tip.nl zelf hele zinnen vertalen met deze methode.

OPGAVE 1. Bovenstaande formule kun je ook gebruiken om het geheimschrift weer terug te vertalen naar de originele tekst. Alleen de waarden van a en b zijn anders. Bereken deze waarden.

2. Niet alle waarden voor a zijn bruikbaar. Neem je bijvoorbeeld $a = 4$ en $b = 13$, dan wordt het bovenstaande zinnetje vertaald als: VFLPNLRDNH TH DDN LTXZHVPDTHL TRRD XRNLDDDN. Hier klopt iets niet: zowel de 'e' als de 'r' worden afgebeeld op 'd'. Daardoor is terugvertalen naar de originele tekst niet meer mogelijk. Aan welke voorwaarde moet de keuze voor a voldoen?

De code kraken

Neem eens aan dat je de volgende tekst hebt onderschept: RLB TTGFKTTGIN OTKNAITX GHNK GTTA INUNGKNA. IN YWLNBNG RPGGNG GHNK ZLAING YNOLAXI. Je wilt natuurlijk graag weten wat hier staat. Je bent getipt dat hier gebruik is gemaakt van de formule $F(x) = ax + b$. Voor het ontcijferen passen we eerst *frequentieanalyse* toe; dat betekent dat we voor elke letter van het alfabet nagaan hoe vaak deze letter voorkomt. Dat kan met de hand, maar ook met het programma Numbers (zie www.crypto.club.tip.nl). Het resultaat is:

```
Total number of letters: 69
A:  5   7.25%     B:  2   2.90%
C:  0   0.00%     D:  0   0.00%
E:  0   0.00%     F:  1   1.45%
G: 11  15.94%     H:  2   2.90%
I:  6   8.70%     J:  0   0.00%
K:  5   7.25%     L:  4   5.80%
M:  0   0.00%     N: 13  18.84%
O:  2   2.90%     P:  1   1.45%
Q:  0   0.00%     R:  2   2.90%
S:  0   0.00%     T:  8  11.59%
U:  1   1.45%     V:  0   0.00%
W:  1   1.45%     X:  2   2.90%
Y:  2   2.90%     Z:  1   1.45%
```

We zien dat de letters N en G het meeste voorkomen. We gaan er daarom van uit dat N de letter E was en G de letter N. Nu moeten we op zoek gaan naar getallen a en b zodat geldt:

$$4 \equiv 13a + b \pmod{26}$$
$$13 \equiv 6a + b \pmod{26}$$

Dit stelsel gaan we oplossen. Wanneer we de tweede vergelijking van de eerste aftrekken, dan krijgen we $17 \equiv 7a \mod 26$. Modulo 26 is de inverse van vermenigvuldigen met 7 vermenigvuldigen met 15 (zie het decembernummer). Dus $a = 17 \times 15$, en we zien $a \equiv 21 \mod 26$. Als we dit in de eerste vergelijking invullen dan krijgen we: $4 = 21 \times 13 + b \mod 26$ en $b = 17$. Met de formule $G(x) = 21x + 17$ kunnen we dus het geheimschrift ontcijferen. Er blijkt te staan: "Kom aanstaande zaterdag niet naar Deventer. De bloemen kunnen niet worden bezorgd."

OPGAVE 3. Wat staat hier?
OXE CXBFX YJW RXOXLZHDOKLUEXW LH AJE WLXZJWA JWAXKH MX FJW CXMXW!

4. En wat staat hier?
FH SFKA EBQ EBIBJXXI KFBQ WL IBRH

AXQ GB AFQ HXK IBWBK.

5. Tenslotte, wat staat hier?

XJ IXZ XJ IXZ DBY QXQ OXZY IXZY ZO GZ JCZRM XF DXY.

Een cryptografische aanval

We hebben gezien dat je met een formule de letters van het alfabet door elkaar kan husselen. We gaan nu proberen dergelijke geheimschriften op een systematische manier te kraken – dit heet een cryptografische aanval. We beginnen met het tellen van de mogelijkheden. Hoeveel formules $ax + b$ zijn er? Voor a mag je alle getallen nemen die geen delers met 26 gemeen hebben. Voor b mag je alle getallen van 0 tot en met 25 kiezen. Dat levert in totaal $12 \times 26 = 312$ verschillende formules $ax + b$. Als je weet dat een tekst met deze methode vercijferd is, dan kun je met een computer al deze ontcijferformules uitproberen. Je krijgt dan 312 verschillende zinnen. Bijna alle zinnen zijn abracadabra, maar precies één ervan is de oorspronkelijke tekst. Je moet dus zo'n driehonderd zinnen controleren: staat er onzin of een geheime boodschap? Deze controle kun je zelf doen, maar er bestaan ook programma's die dit voor je doen.

Voor een computer is driehonderd mogelijkheden niet veel. Je kunt de methode verbeteren door 29 letters te gebruiken in plaats van 26. Dit kan bijvoorbeeld door aan het alfabet de spatie, de komma en de punt toe te voegen. Omdat 29 een priemgetal is, zijn er dan $(29 - 1) \times 29 = 812$ mogelijkheden. Erg veel schiet je daar niet mee op: het aantal mogelijkheden wordt maar drie keer zo groot.

Een betere methode?

Een echte verbetering ontstaat als je in plaats van losse letters groepjes van twee letters samen neemt. Er zijn dan in totaal $26^2 = 676$ verschillende lettercombinaties. Net als boven kun je dan weer formules $ax + b$ bedenken om deze lettercombinaties te vercijferen. Je kunt dan modulo 676 rekenen of nog beter modulo 677 (een priemgetal). Bij deze laatste mogelijkheid zijn er $677 \times (677 - 1) = 457652$ verschillende waarden voor a en b.

Zoveel mogelijkheden lijkt veel, maar echt veilig is deze verbeterde vercijferingsmethode niet. Ook voor paren van letters kun je een frequentieanalyse doen en er bestaan tabellen waarmee je aan de slag kunt. Maar dat is niet de enige zwakte van deze methode. Laten we eens aannemen dat je een tekst hebt onderschept van 400 lettertekens en dat het vertalen van één letterteken eenmiljoenste seconde duurt. Dan moeten we 457652×400 keer een letter vertalen. Dit duurt ongeveer drie minuten! We zien dus dat onze methode niet echt veilig is: met een computer kun je in zeer korte tijd alle mogelijkheden uitproberen. Deze manier om een geheimschrift te kraken wordt de *brute kracht*-methode genoemd (brute force in het Engels): 'bruut' omdat deze methode niet erg elegant is en 'kracht' omdat gebruik gemaakt wordt van de rekenkracht van de moderne computer. ◢