scienceinfuse

ANTENNE DE FORMATION ET DE PROMOTION DU SECTEUR SCIENCES & TECHNOLOGIES

DOSSIER ENSEIGNANT

π

MATHS

Une introduction à la cryptographie et au système RSA

UCL **Scienceinfuse** - Antenne de formation et de promotion du secteur sciences & technologies
rue des Wallons 72 L6.02.01 - 1348 Louvain-la-Neuve

La **cryptographie** est utilisée pour dissimuler des messages aux yeux de certains utilisateurs. Elle désigne l'ensemble des techniques permettant de chiffrer (ou coder) des messages, c'est-à-dire permettant de les rendre incompréhensibles.

La **cryptanalyse** est la reconstruction d'un message chiffré en clair (ou décodage) à l'aide de méthodes mathématiques.

La **cryptologie** est la science qui étudie les aspects scientifiques de ces techniques, c'est-à-dire qu'elle englobe la cryptographie et la cryptanalyse. La cryptologie est essentiellement basée sur l'arithmétique.

Exemple :

"OPVTBUUBRVPOTEFNBJO"

signifie

"Nous attaquons demain"

Pour comprendre ce message il faut connaître la clé qui a servi à le coder.

Le fait de coder un message pour le rendre secret s'appelle **chiffrement**. La méthode inverse, qui consiste à retrouver le message original, est appelée **déchiffrement**.
Nous allons voir ici deux types de chiffrement. Il en existe beaucoup d'autres.

1 Chiffrement par décalage

Jules César ne faisait pas confiance à ses messagers lorsqu'il envoyait des messages à ses généraux. Il chiffrait ses messages en remplaçant tous les "A" par des "D", les "B" par des "E" et ainsi de suite. Seule la personne connaissant la clé correspondant au nombre de caractères de décalage (ici 3) pouvait déchiffrer ses messages.

Activité 1 : En utilisant les bandelettes et le code secret de Jules César, chiffrez le mot "bonjour" et déchiffrez la phrase "frpphqw ydv-wx ?"

Placer les élèves par deux et donner à chaque groupe une grande et une petite bandelette.
Il faut mettre la grande bandelette à gauche et la petite à droite et aligner le "A" du milieu de la grande bandelette avec le "D" de la petite bandelette. Le codage se fait de gauche à droite et le décodage de droite à gauche.

Si on décale de 3 l'alphabet	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
on obtient	D E F G H I J K L M N O P Q R S T U V W X Y Z A B C
et donc	B O N J O U R
devient	E R Q M R X U
et	F R P P H Q W Y D V – W X ?
signifie	C O M M E N T V A S – T U ?

L'avantage de ce procédé est qu'il est facile à utiliser pour coder les messages.
Son inconvénient est que les messages codés par ce procédé sont faciles à décoder. En effet,

- les mêmes lettres sont codées par des mêmes lettres (par exemple les deux "O" dans "bonjour") ;
- même si on ne connaît pas la clé, il n'y a que 25 possibilités de décaler les lettres de l'alphabet.

Un autre procédé de chiffrement consiste à retourner l'alphabet puis décaler les lettres.

Activité 2 : A l'aide des bandelettes, utilisez ce procédé avec un décalage de 5 pour chiffrer le mot "bonjour".

Placer les élèves par deux et donner à chaque groupe une grande et une petite bandelette.
Il faut mettre la grande bandelette à gauche et la petite à droite et retourner la petite bandelette. Aligner le "A" du milieu de la grande bandelette avec le "U" de la petite bandelette.
Le codage se fait de gauche à droite.

Si on renverse l'aphabet	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
on obtient	Z Y X W V U T S R Q P O N M L K J I H G F E D C B A
puis on décale de 5	U T S R Q P O N M L K J I H G F E D C B A Z Y X W V
et donc	B O N J O U R
devient	T G H L G A D

Ce procédé est facile à utiliser pour coder les messages. De plus, les messages codés par ce procédé sont plus difficiles à décoder que par le code de Jules César si on ne pense pas à retourner d'abord l'alphabet.
Cependant, les messages codés par ce procédé sont quand même faciles à décoder. En effet,

- les mêmes lettres sont codées par des mêmes lettres (par exemple les deux "O" dans "bonjour") ;
- si on sait qu'il faut d'abord retourner l'alphabet, il n'y a que 26 possibilités à essayer.

2 Chiffrement cyclique

Pour que les messages soient plus difficiles à décoder on peut utiliser le **chiffrement cyclique**.
Il s'agit de décaler les lettres de l'alphabet mais en changeant le décalage à chaque lettre.
Pour cela, il faut choisir un **mot-clé** qui nous indiquera le décalage à effectuer.
Par exemple, le mot-clé "SCIENCES" signifie que pour décoder la première lettre on aligne "A" avec "S", pour décoder la deuxième lettre on aligne "A" avec "C", pour décoder la troisième lettre on aligne "A" avec "I" et ainsi de suite. Après la huitième lettre (fin du mot SCIENCES), on recommence "A" avec "S",...

Activité 3 : A l'aide du mot-clé "SCIENCES", chiffrez le mot "bonjour".

Cette activité peut se faire à l'aide des bandelettes, avec un disque de chiffrement ou la Table de Vigenère.

Avec les bandelettes : il faut mettre la grande bandelette à gauche et la petite à droite et aligner le "A" du milieu de la grande bandelette avec le "S" de la petite bandelette pour la première lettre, puis aligner le "A" du milieu de la grande bandelette avec le "C" de la petite bandelette pour la deuxième lettre et ainsi de suite. Le codage se fait de gauche à droite.

Avec le disque de chiffrement : les lettres à coder se trouvent sur le disque de couleur. A chaque lettre, aligner le "A" du disque de couleur avec la lettre correspondante du mot-clé sur le disque blanc. Le mot codé est construit en utilisant les lettres du disque blanc.

Avec la Table de Vigenère : pour chaque lettre en clair, on sélectionne la colonne correspondante et pour une lettre de la clé on sélectionne la ligne adéquate, puis au croisement de la ligne et la colonne on trouve la lettre chiffrée. La lettre de la clé est à prendre dans l'ordre dans laquelle elle se présente et on répète la clé en boucle autant que nécessaire.

On place le "S" en face du "A" et donc "B" devient "T".
On place le "C" en face du "A" et donc le "O" devient "Q".
On place le "I" en face du "A" et donc le "N" devient "V".
On place le "E" en face du "A" et donc le "J" devient "N".
On place le "N" en face du "A" et donc le "O" devient "B".
On place le "C" en face du "A" et donc le "U" devient "W".
On place le "E" en face du "A" et donc le "R" devient "V".

Le mot "BONJOUR" est donc codé par "TQVNBWV".

Les avantages de ce procédé sont les suivants :

- il est facile à utiliser pour coder les messages ;
- les mêmes lettres sont codées par des lettres différentes (par exemple les deux "O" dans "bonjour") ;
- des lettres différentes sont codées par les mêmes lettres (par exemple "N" et "R" sont toutes les deux codées par "V" dans "bonjour").

Les messages codés par ce procédé sont très difficiles à décoder si on ne connaît pas le mot-clé.

3 Cryptographie à clé publique - RSA

Le chiffrement se fait généralement à l'aide d'une clé de chiffrement, pour le déchiffrement on utilise une clé de déchiffrement. Il existe deux types de clés :

- Les **clés symétriques** : la même clé est utilisée pour le chiffrement et pour le déchiffrement. On parle alors de chiffrement symétrique ou de chiffrement à clé secrète. C'est ce type de clé qui a été utilisé dans les activités ci-dessus.
- Les **clés asymétriques** : une clé différente est utilisée pour le chiffrement et pour le déchiffrement. On parle alors de chiffrement asymétrique ou de chiffrement à clé publique.

Quel type de clés utiliser ?

D'un point de vue informatique, un programme de chiffrement/déchiffrement à clé secrète est très rapide tandis qu'un programme à clé publique peut être beaucoup plus lent (car lourd en calcul). Aussi, plutôt que d'envoyer des messages entiers avec un système à clé publique (qui a l'avantage de ne pas devoir se mettre d'accord au préalable sur une clé commune), on l'utilise juste une fois pour se mettre d'accord sur un mot-clé commun. Ensuite on s'envoie les messages en utilisant un système à clé secrète en utilisant ce mot-clé.
L'idée est donc de commencer par utiliser des clés asymétriques pour s'échanger le mot-clé. Une fois que l'expéditeur et le destinataire du message connaissent le mot-clé, ils l'utilisent comme clé symétrique pour coder et décoder des messages.

Comment procéder à l'échange du mot-clé de manière sécurisée sans aucun dispositif de sécurité ?

La **cryptographie à clé publique** utilise deux clés pour le cryptage : une clé publique pour crypter les données et une clé privée pour les décrypter. On peut ainsi publier la clé publique tout en conservant la clé privée secrète. D'un point de vue informatique, il est impossible en un temps raisonnable de deviner la clé privée à partir de la clé publique. Un utilisateur qui possède une clé publique peut donc crypter des informations mais est dans l'impossibilité de les décrypter. Seule la personne disposant de la clé privée correspondante peut décrypter ces informations.

Cette méthode présente un très gros avantage : elle permet d'échanger des messages de manière sécurisée **sans aucun dispositif de sécurité**. L'expéditeur et le destinataire n'ont plus besoin de partager des clés secrètes par une voie de transmission sécurisée car les communications impliquent uniquement l'utilisation de clés publiques. Aucune clé privée n'est transmise ou partagée.

Nous allons voir ici un processus de cryptographie à clé publique appelé **RSA** (d'après le nom de ses inventeurs Ronald Rivest, Adi Shamir et Leonard Adleman en 1977). Il est très utilisé dans le commerce électronique et pour échanger des données confidentielles sur Internet.
Le chiffrement RSA est asymétrique. Il utilise deux clés (des nombres entiers) : une clé publique pour chiffrer et une clé privée pour déchiffrer les données. Une personne A (souvent nommée Alice) souhaite que B (souvent appelé Bob) lui envoie des données confidentielles. La personne A crée les deux clés, rend la clé publique accessible et conserve la clé privée. La clé publique est utilisée par son correspondant B pour chiffrer les données qui seront envoyées. La clé privée permettra à A de déchiffrer ces données.
Ce procédé est souvent utilisé pour transmettre le mot-clé qui permettra alors de poursuivre l'échange de message par chiffrement cyclique comme vu ci-dessus : Bob envoie à Alice une clé de chiffrement symétrique (le mot-clé) qui peut ensuite être utilisée par Alice et Bob pour échanger des informations.

Voyons ce processus plus en détails : supposons que B souhaite envoyer le mot-clé à A.

- A construit une clé privée et une clé publique. Elle envoie la clé publique à B sans plus de précaution.
- B se sert de cette clé pour crypter le mot-clé. Il envoie le mot-clé crypté à A.
- Grâce à la clé privée, A parvient à décoder le mot-clé.

Etape 1 : Préparation des clés et envoi de la clé publique

- A choisit deux nombres premiers p et q tenus secrets.
  A construit $n = p \cdot q$, $n > 26$.
  A calcule $f = (p-1)(q-1)$.
  A choisit e tel que $e < f$ et e premier avec f.
  La clé publique est (e, n).
- A calcule d tel que $e.d = 1 \mod f$.
  La clé privée de A est d.
- A envoie la clé publique (e, n) à B.

Etape 2 : Codage et envoi du mot-clé

- B transforme le mot-clé en nombres selon la position de ses lettres dans l'alphabet.
  Le mot-clé est donc transformé en nombres x.
  Remarque : en pratique la transformation du texte en nombres est un peu plus complexe (il faut éviter que 2 lettres identiques soient envoyées sur 2 nombres identiques).
- B code chaque nombre x du mot-clé à l'aide de la clé publique : $y = x^e \mod n$.
- B envoie les nombres y à A.

Etape 3 : Décodage du mot-clé

- A décode les nombres y à l'aide de sa clé privée : $z = y^d \mod n$.
- Or $z = x$ et donc A peut reconstituer le mot-clé.

Activité 4 : Utilisez le système RSA pour échanger le mot-clé "SCIENCES".

B souhaite envoyer le mot-clé "SCIENCES" à A.

Etape 1 : Préparation des clés et envoi de la clé publique

- A choisit deux nombres premiers p et q tenus secrets, par exemple $p = 3$ et $q = 11$.
  A construit $n = p \cdot q$, donc $n = 33$.
  A calcule $f = (p-1)(q-1)$, ici $f = 20$.
  A choisit e tel que $e < f$ et e premier avec f, par exemple $e = 7$.
  La clé publique est $(e, n) = (7, 33)$.
- A calcule d tel que $e \cdot d = 1 \mod f$: $7 \cdot d = 1 \mod 20$ donc $d = 3$.
  La clé privée de A est $d = 3$.
- A envoie la clé publique $(7, 33)$ à B.

Etape 2 : Codage et envoi du mot-clé

- B transforme le mot-clé en nombres selon la position de ses lettres dans l'alphabet : SCIENCES = 19 3 9 5 14 3 5 19.
- B code chaque nombre x du mot-clé à l'aide de la clé publique : $y = x^7 \mod 33$:
  $\rightsquigarrow$ 19 devient 13 car

$$\begin{aligned} 19^7 \mod 33 &= 19 \cdot 361 \cdot 361 \cdot 361 \mod 33 \\ &= 19 \cdot 31 \cdot 31 \cdot 31 \mod 33 \\ &= 566029 \mod 33 = (33 \cdot 17152 + 13) \mod 33 = 13 \end{aligned}$$

  $\rightsquigarrow$ 3 devient 9 car

$$3^7 \mod 33 = 2187 \mod 33 = (33 \cdot 66 + 9) \mod 33 = 9$$

$\rightsquigarrow$ 9 devient 15 car

$$\begin{aligned} 9^7 \bmod 33 &= 9 \cdot 81 \cdot 81 \cdot 81 \bmod 33 \\ &= 9 \cdot 15 \cdot 15 \cdot 15 \bmod 33 \\ &= 30375 \bmod 33 = (33 \cdot 920 + 15) \bmod 33 = 15 \end{aligned}$$

$\rightsquigarrow$ 5 devient 14 car

$$\begin{aligned} 5^7 \bmod 33 &= 5 \cdot 125 \cdot 125 \bmod 33 \\ &= 5 \cdot 26 \cdot 26 \bmod 33 \\ &= 3380 \bmod 33 = (33 \cdot 102 + 14) \bmod 33 = 14 \end{aligned}$$

$\rightsquigarrow$ 14 devient 20 car

$$\begin{aligned} 14^7 \bmod 33 &= 14 \cdot 196 \cdot 196 \cdot 196 \bmod 33 \\ &= 14 \cdot 31 \cdot 31 \cdot 31 \bmod 33 \\ &= 417074 \bmod 33 = (33 \cdot 12638 + 20) \bmod 33 = 20 \end{aligned}$$

Donc

$$19\ 3\ 9\ 5\ 14\ 3\ 5\ 19$$

devient

$$13\ 9\ 15\ 14\ 20\ 9\ 14\ 13.$$

- B envoie à A les nombres 13 9 15 14 20 9 14 13.

Etape 3 : Décodage du mot-clé

- A décode les nombres y à l'aide de sa clé privée : $z = y^3 \bmod 33$:
  $\rightsquigarrow$ 13 devient 19 car

$$13^3 \bmod 33 = 2197 \bmod 33 = (33 \cdot 66 + 19) \bmod 33 = 19$$

  $\rightsquigarrow$ 9 devient 3 car

$$9^3 \bmod 33 = 729 \bmod 33 = (33 \cdot 22 + 3) \bmod 33 = 3$$

  $\rightsquigarrow$ 15 devient 9 car

$$15^3 \bmod 33 = 3375 \bmod 33 = (33 \cdot 102 + 9) \bmod 33 = 9$$

  $\rightsquigarrow$ 14 devient 5 car

$$14^3 \bmod 33 = 2744 \bmod 33 = (33 \cdot 83 + 5) \bmod 33 = 5$$

  $\rightsquigarrow$ 20 devient 14 car

$$20^3 \bmod 33 = 8000 \bmod 33 = (33 \cdot 242 + 14) \bmod 33 = 14$$

- Or $z = x \bmod 33$ et donc A peut reconstituer le mot-clé :

$$13\ 9\ 15\ 14\ 20\ 9\ 14\ 13 \text{ devient } 19\ 3\ 9\ 5\ 14\ 3\ 5\ 19 = \text{SCIENCES}.$$

Un système de chiffrement à clé publique fonctionnera bien à condition qu'il soit mathématiquement impossible de reconstituer la clé privée à partir de la clé publique en un temps raisonnable, c'est-à-dire que les moyens de calcul disponibles et les méthodes connues au moment de l'échange ne le permettent pas.

Dans le cas du RSA, on choisira les nombres p et q très grands de sorte qu'il soit très difficile (c'est-à-dire impossible en un temps raisonnable) de les retrouver en ne connaissant que leur produit $n = p \cdot q$. Par exemple, pour $n = 2351537$, qui pourra retrouver les nombres premiers p et q tels que $p \cdot q = n$? Ainsi, un espion qui connaît n ne pourra pas trouver p et q et donc ne pourra pas deviner f ni d (la clé privée).
Notons que si n est petit alors l'espion peut essayer différentes valeurs pour d. **Les nombres utilisés dans ce processus doivent donc être très grands**.

Activité 5 : Envoyez un message secret à votre ami. Pour cela, commencez par échanger un mot-clé en utilisant le système RSA, puis utilisez le disque de chiffrement ou les bandelettes pour coder votre message et décoder le message reçu de votre ami.

Pour cette activité, placer les élèves par deux. Chaque élève écrit un petit mot qu'il veut envoyer à son ami. Ensuite les élèves échangent un mot-clé comme décrit dans le système RSA. Une fois le mot-clé échangé, chaque élève code son message à l'aide du mot-clé et l'envoie à son ami qui le décode à l'aide du disque de chiffrement ou des bandelettes.