Théorie de l'information, théorie des nombres et mathématique discrète

Razafimahefa F. O. [1], Randriamitantsoa P.A. [2], Randriamitantsoa A. A[3]

Laboratoire de Recherche en Télécommunication, Automatique, Signal et Images (LR-TASI)

École Doctorale en Sciences et Techniques de l'Ingénierie et de l'Innovation (ED-STII)

Équipe d'Accueil Doctorale Télécommunication, Automatique, Signal et Images (EAD-TASI)

Université d'Antananarivo

BP 1500, Ankatso – Antananarivo 101 – Madagascar

[1]ombana.razafimahefa@gmail.com , [2]rpauguste@gmail.com, [3]andriau23@gmail.com

Résumé

Dans cet article, nous allons nous étaler sur des notions mathématiques indispensables à la cryptographie. Il est divisé en deux grandes parties. Celle relative à la théorie de l'information dans un premier temps. La seconde partie en est la théorie des nombres et la mathématique discrète.

Dans la théorie de l'information, nous allons voir entre autres les deux principaux théorèmes de Shannon : celui du codage source et du codage canal. Le premier stipule que le taux du code ne devrait être inférieur à l'entropie de la source, sinon l'on perdrait l'information. Le second affirme que pour un niveau de bruit donné, il existe un débit suivant lequel, l'on pourrait transmettre un message dans le canal.

Dans la seconde partie qu'est la théorie des nombres et la mathématique discrète, nous allons remettre en exergue des notions d'arithmétiques modulaires comme quoi les propriétés (commutativité, associativité, distributivité, identité et inverse) que nous connaissons avec les opérations standards ($+,-$ et $\times$) tiennent également dans l'ensemble des restes de la division par N, l'ensemble $\mathbb{Z}_N$. Nous allons également développer la méthode de Miller-Rabin qui est une approche probabiliste en termes de confirmation ou non si un nombre n donné est premier.

Mots clés : *Théories de l'information, théorie des nombres, mathématique discrète.*

Abstract

In this article, we will expand on the mathematical concepts essential to cryptography. It is divided into two main parts. The first relating to information theory. The second part is number theory and discrete mathematics.

In information theory, we will see among other things the two main theorems of Shannon: that of source coding and channel coding. The first

stipulates that the code rate should not be lower than the entropy of the source, otherwise the information will be lost. The second asserts that for a given noise level, there is a rate at which a message could be transmitted in the channel.

In the second part, which is number theory and discrete mathematics, we are going to highlight notions of modular arithmetic such as the properties (commutativity, associativity, distributivity, identity and inverse) that we know with standard operations ($+, -$ and $\times$) also hold in the set of the remains of the division by N, the set $\mathbb{Z}_N$. We will also develop the Miller-Rabin method which is a probabilistic approach in terms of confirmation or not if a given number n is prime.

Keywords: *Information theories, source coding, channel coding, confusion, diffusion, complete transformation, avalanche effect, Hamming distance and weight, Kerckhoffs principle, and/or Shannon's maxim, binary mathematics, number theory, discrete mathematics, divisibility, division algorithm, GCD or Greatest Common Divisor, modular arithmetic, Euclid algorithm, extended Euclid algorithm, Chinese remainder theorem, factorization, prime numbers, Fermat theorem, Miller-Rabin algorithm.*

1 Théorie de l'information

1.1 Premier théorème de Claude Shannon ou théorème du codage source

Le théorème stipule qu'il est impossible de compresser une chaine de variables aléatoires indépendantes, identiquement distribuées (i.i.d.) – la longueur de la chaîne tendant donc vers l'infini – de telle sorte que le taux du code (la longueur moyenne des codes des variables) soit inférieur à l'entropie de la source, sans avoir la certitude que l'on perde de l'information. L'on peut cependant avoir une compression avec un taux de code proche de ladite entropie. [2], [3]

1.2 Deuxième théorème ou théorème du codage canal

Ce théorème statue qu'il est possible de transmettre des données numériques sur un canal, même bruité, presque sans erreur à un débit maximum calculable.

La limite ou capacité de Shannon d'un canal est le débit maximal de transfert d'information sur ce canal pour un certain niveau de bruit donné. [2], [4]

La formulation mathématique du théorème est attribuée par la littérature au binôme *Shannon-Hartley* et est donnée par ce qui suit :

$$C = B \log\left(1 + \frac{S}{N}\right) \qquad (1.01)$$

Où :

- C est la capacité du canal en bits par seconde (bits/s) ;

- B est la bande passante du canal en hertz (Hz) ;

- S/N est le rapport signal sur bruit (*SNR* ou *Signal to Noise Ratio* en anglais) ;

1.3 Entropie d'une information

L'entropie d'une information se définit comme étant la mesure de la quantité d'information contenue dans un message. Cette définition équivaut au fait que l'entropie est également la mesure de l'incertain dans un message. [2], [5], [6]

Elle est donnée par la formule (1.02) qui suit.

$$H(X) = -\sum_{k=1}^{K} p_k \log p_k \qquad (1.02)$$

Où :

X dénote une variable aléatoire i.i.d. et $H(X)$ l'entropie de X ;

K le nombre de lettres possibles de l'alphabet dans lequel sont tirés les lettres du message composant X ;

p_k la probabilité d'apparition de chaque lettre a_k de l'alphabet ;

1.4 Confusion et diffusion

Confusion et diffusion sont deux propriétés de méthode cryptographique énumérées par Claude Shannon. [2]

La confusion se veut d'être une propriété selon laquelle la relation entre le texte en clair et la clé de chiffrement doit être difficilement décelable. Une lettre chiffrée devrait donc dépendre de plusieurs lettres de la clé de chiffrement à la fois. Par opposition au fait de dépendre uniquement d'une seule lettre de la clé de chiffrement.

La diffusion quant à elle est une propriété qui garantit que toute redondance que l'on pourrait trouver dans le texte en clair ne le soit plus dans le texte chiffré. Soit en d'autres termes, elle (la redondance) est dissipée d'une manière convenable à une partie du texte chiffré. [2]

1.5 Transformation complète et effet d'avalanche

Une transformation est dite complète si chacune des lettres du texte chiffré dépend de toutes les lettres du texte en clair. Ainsi, s'il est possible de trouver une version simplifiée de ladite transformation pour chacune des lettres chiffrées, alors elle dépendrait en son entrée de toutes les lettres en claires. [6]

L'avalanche est une propriété de méthode cryptographique qui se veut qu'une petite modification dans le texte en claire ait une

répercussion importante dans le texte chiffré, comme une avalanche donc. Cette propriété a une similitude à celle qu'est la diffusion de Claude Shannon, l'avalanche est cette fois-ci attribuée par la littérature à *Horst Feistel*. Ce dernier était un chercheur américano-allemand en cryptographie, recherches qui ont culminé avec l'apparition de la *Data Encryption Standard* ou *DES* dans les années 1970. [6]

Le critère d'avalanche strict est une formalisation de l'effet d'avalanche. Il est satisfait si le fait de modifier un seul bit du texte en clair conduit à une modification de tous les bits de sortie du message chiffré avec une probabilité de 0.5, soit la moitié des bits chiffrés change en d'autres termes. Le concept de critère d'avalanche strict a été introduit par *Webster* et *Tavares* en 1985. [6]

1.6 Distance et poids de Hamming

La distance de Hamming est donnée par le nombre de caractères qui diffèrent entre deux textes donnés. Elle est dénotée par $h(x, y)$. La distance de Hamming est en d'autres termes le nombre de substitutions nécessaires pour aller d'un texte à un autre. [2]

Pour notre ère, qui est caractérisée par la cryptographie moderne, l'information représentée sous forme de bits. La distance de Hamming est donc donnée par le nombre de bits 1 résultant de l'opération eXclusive OR ou XOR entre les deux textes x et y.

$$h(x, y) = x \oplus y \qquad (1.03)$$

Où $\oplus$ représente l'opération XOR.

Le poids de Hamming est un peu identique au concept de distance de Hamming. Il s'agit de compter le nombre de bits égal à 1 dans un texte donné. Elle équivaut donc à calculer la distance de Hamming du texte à un second autre dont chacun des bits est égal à 0. [2]

1.7 Principe de Kerckhoffs et/ou maxime de Shannon

Le principe de Kerckhoffs se veut que « la sécurité d'un chiffre (méthode cryptographique) dépende uniquement du fait que la clé de chiffrement est tenue secrète et non du fait que les détails de l'algorithme lui-même soient tenus secrets ». Shannon, de par sa maxime a évoqué la formulation qui suit : « Nous devrions construire notre chiffre (méthode cryptographique toujours) sous la connaissance que l'ennemi pourrait être en possession de ladite algorithme ». [2]

Le principe de Kerckhoffs et la maxime de Shannon sont donc d'accords sur un point : seul la clé doit être tenue secret. Le fait de connaitre donc que quelqu'un ait utilisé du SHA-256 bits, du Blowfish, du Serpent ou quoi que ce soit d'autres comme type d'algorithme, ne met en rien en périls la sécurité de son message chiffré tant qu'il garde bien secrète sa clé de chiffrement.

1.8 Mathématique binaire

La mathématique binaire se définit comme étant la mathématique qui utilise la base 2 au lieu de la base ordinaire 10. La base 10 est plus connue par la plupart des personnes lambda suite au fait qu'il est plus facile de compter 10 par 10, le nombre de doigts et d'orteils que nous avons aidant sûrement.

Les opérations binaires nécessaires à la cryptographie sont principalement le OR, le AND et le XOR. L'on pourrait également ajouter l'opération NOT. Cette dernière, contrairement au OR, AND and XOR, requiert uniquement une entrée, si deux sont nécessaires pour les 3 premières opérations citées. [2]

Le résultat du OR est égal à 1 si au moins un de ses deux entrées est égale à 1. 0 sinon.

a	b	a OR b
0	0	0
0	1	1
1	0	1
1	1	1

Tableau 1.01 : *Opération binaire OR.*

Le résultat du AND est égal à 1 si toutes ses deux entrées sont égales à 1. 0 sinon.

a	b	a AND b
0	0	0
0	1	0
1	0	0
1	1	1

Tableau 1.02 : *Opération binaire OR.*

Le résultat du XOR est égal à 1 ses deux entrées sont différentes. 0 sinon.

a	b	a XOR b
0	0	0
0	1	1
1	0	1
1	1	0

Tableau 1.03 : *Opération binaire OR.*

L'opération XOR est très importante dans le domaine de la cryptographie. Ceci est due au fait que cette opération est non seulement réversible, mais que son inverse aussi n'est autre qu'elle-même. En d'autres termes, si l'on a $a\ XOR\ b\ = c$, alors l'on pourra retrouver a en évaluant $b\ XOR\ c$.

Le NOT consiste à complémenter un bit donné. Si son entrée est 0, la sortie en est 1. Si à l'entrée l'on a 1, sa sortie est 0.

a	NOT a
0	1
1	0

Tableau 1.04 : *Opération binaire NOT.*

2 Théorie des nombres et mathématique discrète

La théorie des nombres se définit comme étant l'étude des nombres entiers positifs. L'étude des nombres premiers et de leur factorisation constitue un aspect important de la théorie des nombres, du moins pour son application dans le domaine de la cryptographie.

Comme avant-goût, l'on pourrait citer par exemple l'usage des nombres premiers (ici 2) dans la constitution du corps de Galois (ou Galois Field an anglais) $GF(2^8)$ sur lequel se base la S-box (boîte à substitution) du chiffre SHARK. SHARK étant un algorithme de cryptographie prédécesseur à l'AES (Advanced Encryption Standard).

2.1 Divisibilité et algorithme de division

L'on dit qu'un nombre entier b divise a si l'on a $a = mb$ où m est également un nombre entier. La notation $b|a$ est majoritairement utilisé pour exprimer le fait que b divise a. b est dans ce cas également appelé *diviseur* de a. [1]

$$b|a \Leftrightarrow a = mb \qquad (2.01)$$

Exemple : Les diviseurs positifs de 24 sont $1, 2, 3, 4, 6, 8, 12$ et 24 lui-même.

Etant donné un nombre entier n, et un autre non négatif entier a, si nous divisons a par n, nous aurons deux entiers q (pour quotient) et r (pour reste) qui obéissent à la formulation : [1]

$$a = qn + r \text{ où } 0 \leq r < n \text{ et } q = \lfloor a/n \rfloor \qquad (2.02)$$

Où l'opérateur $\lfloor x \rfloor$ représente l'arrondissement a l'entier directement inférieur ou égal à x.

Exemple : $a = 11, n = 7$, l'on a $11 = 1 \times 7 + 4$ donc $q = 1$ et $r = 4$.

2.2 Plus Grand Commun Diviseur ou PGCD

Le *Plus Grand Commun Diviseur* de deux nombres entiers a et b est noté $\text{PGCD}(a, b)$. Comme son nom l'indique, c'est le plus grand nombre entier c qui divise à la fois a et b. [1]

Une définition équivalente en est, c est le PGCD de a et b si :

- c divise a et b ;

- Tout diviseur de a et b divise c ;

Exemple : $PGCD(24,60) = 12$.

a et b sont premiers entre eux si $\text{PGCD}(a, b) = 1$.

2.3 Arithmétique modulaire

Si a et n sont deux entiers, alors $a \bmod n$ se défini comme étant le reste de la division de a par n. L'on a : [1]

$$a = q \times n + r \text{ avec } q = \lfloor a/n \rfloor \text{ et que } 0 \leq r < n \qquad (2.03)$$
$$a = \lfloor a/n \rfloor \times n + (a \bmod n)$$

Où $\lfloor x \rfloor$ représente le nombre entier immédiatement inférieur ou égal à x.

Exemple : $11 \bmod 7 = 4$. Etant donné que $11 = 1 \times 7 + 4$.

Deux nombre a et b sont dit congruent modulo n si $a \bmod n = b \bmod n$. Cette propriété est dénoté par $a \equiv b \pmod n$.

$$a \equiv b \pmod n \Leftrightarrow a \bmod n = b \bmod n \quad (2.04)$$

Exemple : $73 \equiv 4 \pmod{23}$

Rappelons que d'après la formule (2.03), l'opération modulo associe à tout entier a un nombre entier de l'ensemble réduit $\{0,1,\dots,n-1\}$. La question se pose alors s'il serait possible d'effectuer des opérations dans cette ensemble. Et il s'avère que la réponse à cette question est oui. Elles (lesdites opérations) se définissent comme suit : [1]

- $[(a \bmod n) + (b \bmod n)] \bmod n = (a + b) \bmod n$;

- $[(a \bmod n) - (b \bmod n)] \bmod n = (a - b) \bmod n$;

- $[[(a \bmod n) \times (b \bmod n)] \bmod n = (a \times b) \bmod n$;

Nous allons énumérer dans le tableau ci-dessous (**Tableau 2.01**) quelques propriétés relatives à l'arithmétique modulaire. [1]

Exemple : Pour l'additive inverse dans $\mathbb{Z}_8$, 2 est l'additive inverse de 6. De même 5 est l'additive inverse de 3. En effet, l'on a respectivement $(2+6) \bmod 8 = 8 \bmod 8 = 0$ et $(5+3) \bmod 8 = 8 \bmod 8 = 0$;

Pour la multiplicative inverse, dans le même ensemble $\mathbb{Z}_8$, la multiplicative inverse de 3 est 3. En effet, $(3 \times 3) \bmod 8 = 9 \bmod 8 = 1$. Rappelons que 3 a une multiplicative inverse dans $\mathbb{Z}_8$ car 3 et 8 sont premiers entre eux. Ainsi, 2 par exemple n'a pas de multiplicative inverse dans $\mathbb{Z}_8$ car 2 et 8 ne sont pas premiers entre eux. En effet $\mathrm{PGCD}(2,8) = 2$.

Propriétés	**Expression**
Loi commutatives	$(w+x) \bmod n = (x+w) \bmod n$; $(w \times x) \bmod n = (x \times w) \bmod n$;
Loi associative	$[(w+x)+y] \bmod n = [w+(x+y)] \bmod n$; $[(w \times x) \times y] \bmod n = [w \times (x \times y)] \bmod n$;
Loi distributive	$[w \times (x+y)] \bmod n = [(w \times x) + (w \times y)] \bmod n$;
L'identité	$(0+w) \bmod n = w \bmod n$; $(1 \times w) \bmod n = w \bmod n$;
Existence d'une additive inverse	Pour tout $w \in \mathbb{Z}_n$, $\exists\, z$ tel que $w + z = 0 \bmod n$;

Existence d'une multiplicative inverse, si $\boldsymbol{w}$ et $\boldsymbol{n}$ sont premiers entre eux	Pour tout $w \in \mathbb{Z}_n$, et que w et n sont premiers entre eux, $\exists\, z$, tel que $w \times z \bmod n = 1$

Tableau 2.01 : *Quelques propriétés de l'arithmétique modulaire.*

2.4 Algorithme d'Euclide et Algorithme d'Euclide Etendu

L'algorithme d'Euclide est un algorithme permettant de calculer le PGCD de deux nombres a et b donnés. [1]

Rechercher $\mathrm{PGCD}(a,b)$ revient à rechercher le $\mathrm{PGCD}(b,r_1)$. Ou encore d'une manière plus générale, à rechercher le $\mathrm{PGCD}(r_i,r_{i+1})$

$$\mathrm{PGCD}(a,b) = \mathrm{PGCD}(r_n,r_{n+1}) \text{ où } a \leq b, n \geq 0, r_0 = b\,; \quad (2.05)$$

$$\mathrm{PGCD}(a,b) = \mathrm{PGCD}(r_n,r_{n+1}) = r_n \Leftrightarrow r_{n+1} = 0\,; \quad (2.06)$$

Exemple : Recherchons le $\mathrm{PGCD}(24,60)$.

L'on a donc $a = 24$ et $b = 60$. Comme $a < b$, l'on a à intervertir a et b. Donc nous allons continuer avec $a = 60$ et $b = 24$.

Maintenant, nous avons $60 = 2 \times 24 + 12$. L'on a donc $r = 12$. Comme $r = 12 > 0$, l'on a à continuer l'algorithme donc, cette fois ci en prenant $a = 24$ et $b = 12$.

L'algorithme d'Euclide étendu, en plus de calculer le $\mathrm{PGCD}(a,b)$, calcul également les deux entiers x et y justifiant l'équation ci-dessous : [1]

$$ax + by = \mathrm{PGCD}(a,b) \quad (2.07)$$

Pour ce faire, en plus de calculer les quotients et restes q_n et r_n, l'algorithme d'Euclide étendu calcule également les entiers additionnels x_n et y_n. Les étapes à suivre sont décrites dans le **Tableau 2.02**. L'algorithme s'arrête à l'itération $n+1$ où l'on obtient un reste $r_{n+1} = 0$. Ainsi, les valeurs de r_n, x_n et y_n justifient notre formule (2.07) précédemment citée, soit $ax_n + by_n = r_n = \mathrm{PGCD}(a,b)$.

$\boldsymbol{i}$	$\boldsymbol{q_i}$	$\boldsymbol{r_i}$	$\boldsymbol{x_i}$	$\boldsymbol{y_i}$
$\mathbf{-1}$	–	a	1	0
0	–	b	0	1
1	$\lfloor a/b \rfloor$	$a \bmod b$	$x_{-1} - q_1 x_0$	$y_{-1} - q_1 y_0$
2	$\lfloor b/r_1 \rfloor$	$b \bmod r_1$	$x_0 - q_2 x_1$	$y_0 - q_2 y_1$
3	$\lfloor r_1/r_2 \rfloor$	$r_1 \bmod r_2$	$x_1 - q_3 x_2$	$y_1 - q_3 y_2$

...	...	...	...	...
$\boldsymbol{n}$	$\lfloor r_{n-2}/r_{n-1} \rfloor$	$r_{n-2} \bmod r_{n-1}$	$x_{n-2} - q_n x_{n-1}$	$y_{n-2} - q_n y_{n-1}$
$\boldsymbol{n+1}$	$\lfloor r_{n-1}/r_n \rfloor$	$r_{n-1} \bmod r_n = 0$		

Tableau 2.02 : *Algorithme d'Euclide étendu.*

Exemple : effectuons l'algorithme d'Euclide étendu pour $a = 1212$ et $b = 630$. L'on a donc $n = 3, r_n = r_3 = 6$, $x_n = x_3 = 13$ et $y_n = y_3 = -25$. Soit $13 \times 1212 - 25 \times 630 = 6$.

2.5 Théorème des restes chinois

Le théorème des restes chinois peut s'écrire de différentes manières, mais celle qui nous concerne se détaille comme ce qui va suivre. Soit un entier N tel que : [1]

$$N = \prod_k n_k \text{ avec } 1 \leq k \leq K \qquad (2.08)$$

Où les n_k sont deux à deux premiers entre eux, soit $\forall\, 1 \leq k, l \leq K$ avec $k \neq l$, l'on a $\mathrm{PGCD}(n_k, n_l) = 1$. Nous pouvons représenter tout entier A dans $\mathbb{Z}_N$ par un K-uplet dont chaque éléments appartient à $\mathbb{Z}_{n_k}$.

$$A \leftrightarrow (a_1, \dots, a_K) \qquad (2.09)$$

Où $A \in \mathbb{Z}_N$, $a_k \in \mathbb{Z}_{n_k}$et $a_k = A \bmod n_k, \forall\, 1 \leq k \leq K$. Deux assertions en découlent :

- La correspondance dans l'équation (2.09) est bijective entre $\mathbb{Z}_N$ et l'ensemble produit cartésien $\mathbb{Z}_{n_1} \times \dots \times \mathbb{Z}_{n_K}$. Ainsi, pout tout entier $A \in \mathbb{Z}_N$ il existe un unique K-uplet $(a_1, \dots, a_K) \in \mathbb{Z}_{n_1} \times \dots \times \mathbb{Z}_{n_K}$ qui le représente. Et inversement tout K-uplet $(a_1, \dots, a_K) \in \mathbb{Z}_{n_1} \times \dots \times \mathbb{Z}_{n_K}$ correspond à un unique entier $A \in \mathbb{Z}_N$;

- Tout opération effectuée sur un élément $A \in \mathbb{Z}_N$ peut être effectuée indépendamment sur les K-uplet $(a_1, \dots, a_K) \in \mathbb{Z}_{n_1} \times \dots \times \mathbb{Z}_{n_K}$.

Une des propriétés utiles du théorème des reste chinois est qu'il permet de manipuler un nombre entier très large N en utilisant à sa place ses K-uplet, qui sont des entiers plus petits. Le prérequis en est cependant la nécessité de connaitre la factorisation de N.

Pour la formulation de A en fonction du K-uplet $(a_1, \dots, a_K)$, nous procèderons par sa construction. Soit $N_k = {}^{N}/_{n_k}$ pour $1 \leq k \leq K$. Remarquons que $N_k = n_1 \times \dots \times n_{k-1} \times n_{k+1} \times \dots \times n_K$. Soit $N_l \equiv 0 \pmod{n_k}\ \forall\, l \neq k$. Maintenant soit :

$$m_k = N_k \times (N_k^{-1} \bmod n_k) \equiv 1 \pmod{n_k} \text{ pour } 1 \leq k \leq K \qquad (2.10)$$

L'entier m_k de notre formule (2.10) est bien défini et est unique dû au fait que N_k est premier par rapport à n_k et donc possède une

multiplicative inverse $\bmod n_k$. Maintenant, A peut se calculer comme suit :

$$A \equiv \sum_{k=1}^{K} (a_k m_k) \bmod M \qquad (2.11)$$

Pour montrer que notre formulation de A dans l'équation (2.11) est juste, nous devons montrer que $A \bmod n_k = a_k$ pout tout $1 \leq k \leq K$. Rappelons que $m_l \equiv 0 \pmod{n_k}$ pour tout $l \neq k$. Ce qui revient à dire que $A \bmod n_k \equiv a_k m_k \bmod n_k = a_k$ étant donné que $m_k \bmod n_k = 1$ d'après la définition de m_k dans la formule (2.10).

Exemple : Considérons le cas où $N = 1813 = 37 \times 49$ soient $n_1 = 37$ et $n_2 = 49$, $A = 973$.

Nous avons donc aussi $N_1 = 49$ et $N_2 = 37$. En appliquant l'algorithme d'Euclide étendu, l'on a $N_1^{-1} = 34 \bmod n_1$ et $N_2^{-1} = 4 \bmod n_2$. En prenant le modulo de 973 respectivement par 37 et 49, l'on a la représentation de $A = 973$ qui est $(11, 42)$.

Maintenant, supposons que nous aimerions additionner 678 à 973. L'on a $678 \leftrightarrow (678 \bmod 37, 678 \bmod 49) \equiv (12, 41)$. Il nous suffit d'additionner les 2-uplets un à un et de les réduire. Soit $(11 + 12 \bmod 37, 42 + 41 \bmod 49) = (23, 34)$. Pour voir que notre calcul est correct, recherchons l'équivalent de $(23, 34)$ dans $\mathbb{Z}_{1813}$. Selon le formule (2.11), l'on a $(23, 34) \leftrightarrow (23 \times 49 \times 34 + 34 \times 37 \times 4) \bmod 1813 = 43350 \bmod 1813 = 1651$.

2.6 Factorisation de nombres premiers

Un nombre premier se définit comme étant un nombre positif ayant exactement deux diviseurs : 1 et lui-même. 1 n'est pas considéré comme étant un nombre premier.

Tout nombre entier $a > 1$ peut être écrit d'une manière unique sous la forme [1], [2] :

$$a = p_1^{a_1} \times p_2^{a_2} \times \ldots \times p_K^{a_K} \qquad (2.12)$$

Où $p_1 < p_2 < \cdots < p_K$ sont des nombres premiers et $a_1, a_2, \ldots, a_K$ sont des nombres entiers positifs.

Exemple : $3600 = 24^4 \times 32^2 \times 52^2$; $11011 = 7 \times 11^2 \times 13$;

Soit en d'autres termes si P étant l'ensemble regroupant tous les nombres premiers p_k. L'on aura :

$$a = \prod_{p_k \in P} p_k^{a_k} \qquad (2.13)$$

Où $a_k \geq 0$;

2.7 Théorème de Fermat et quelques propriétés utiles au test de primalité

Le théorème de Fermat statue que, si p est premier et que a est un nombre entier non divisible par p, l'on a : [1], [2]

$$a^{p-1} \equiv 1 \pmod{p} \qquad (2.14)$$

Avant d'entamer les détails du test de primalité de Miller-Rabin, il est nécessaire d'énumérer les quelques notions qui vont suivre. [1], [7]

Pour tout nombre impair n, l'on peut écrire :

$$n - 1 = 2^k q \qquad (2.15)$$

Où k est entier et q est impair.

Une première propriété utile des nombres premiers que nous aurons besoin ici également est que si p est un nombre premier et que a est un nombre entier inférieur à p, alors l'on a :

$$a^2 \bmod p = 1 \Leftrightarrow a \bmod p = 1 \text{ ou } a \bmod p = -1 \equiv p - 1 \qquad (2.16)$$

Où p est premier et $a < p$.

La deuxième propriété d'un nombre premier que nous aurons besoin également est que, si p est un nombre premier plus grand que 2, nous pouvons écrire $p - 1 = 2^k q$ avec k positif et q impair. Soit a un nombre entier tel que $1 < a < p - 1$. Alors, une des deux conditions suivantes est vraie :

- a congrue vers $1 \bmod p$. En d'autres termes, $a^q \bmod p = 1$, ce qui équivaut à $a^q \equiv 1 \pmod{p}$;

- Un des nombre $a^q, a^{2q}, a^{4q}, \dots, a^{2^{k-1}q}$ congruent vers $-1 \bmod p$. En d'autres termes, il y a un nombre j tel que $1 \leq j \leq k$ tel que $a^{2^{j-1}q} \bmod p = -1 \bmod p = p - 1$. Ce qui équivaut à $a^{2^{j-1}q} \equiv -1 \bmod p$.

2.8 Algorithme de Miller-Rabin, un exemple de test de primalité

Nous avons vu dans le paragraphe précédent que si n est premier, alors soit le premier élément de la liste $\left(a^q, a^{2q}, a^{4q}, \dots, a^{2^{k-1}q}, a^{2^k q}\right) \bmod n$ est égal à 1, soit un élément de la liste est égal à $n - 1$; sinon n est un nombre composite. D'un autre côté, si un nombre n vérifie une des deux conditions précédemment énumérées, cela ne signifie pas forcément que n est premier. Considérons par exemple le nombre $n = 2047 = 23 \times 89$. $n - 1 = 2046 = 2 \times 1023$. L'on a $2^{1023} \bmod 2047 = 1$, qui justifie une de nos conditions, alors que 2047 n'est pas premier. [1], [7]

Nous pouvons cependant utiliser le présent algorithme pour diviser le fait qu'un nombre est définitivement composite, ou qu'il pourrait être premier ou non. Dénommons $TEST(n)$ cet algorithme, les étapes correspondant sont comme suit :

$TEST(n)$:

- Trouver les entiers k et q, avec $k > 0$ et q impair tel que $n - 1 = 2^k q$;

- Sélectionner un nombre aléatoire a, tel que $1 <$

$a < n - 1$;

- Si $a^q \bmod n = 1$ alors retourner « n pourrait être premier » ;

- Pour $j = 0$ à $k - 1$ faire si $a^{2^j q} \bmod n = n - 1$ alors retourner « n pourrait être premier » ;

- Retourner « n est composite » ;

Exemple : Essayons de voir si 29 est premier. L'on a $29 - 1 = 28 = 2^2 \times 7 = 2^k q$. Premièrement, essayons $a = 10$. L'on a $10^7 \bmod 29 = 17$. Ce qui n'est ni 1 ni 28. On continue avec $(10^7)^2 \bmod 29 = 28$. Ce qui nous emmène au résultat que 29 pourrait être premier. En essayant encore avec $a = 2$, l'on trouve $2^7 \bmod 29 = 12$ et $(2^7)^2 \bmod 29 = 28$. Ce qui nous emmène, encore une fois, au résultat que 29 pourrait être premier. Si nous réitérons le calcul pour tout $1 < a < 28$, nous obtiendrons toujours le résultat comme quoi n pourrait être premier.

L'algorithme de Miller-Rabin est une méthode à approche probabiliste, comme nous l'avons déjà mentionné un peu plus haut. Il est avancé dans la littérature que cet algorithme a une probabilité inférieure à ¼ de ne pas réussir à détecter qu'un nombre composite l'est (retourne en d'autres termes un résultat comme quoi le nombre pourrait être premier). Répéter l'algorithme de Miller-Rabin t fois revient donc à réduire à $(1/4)^t$ la probabilité de cette erreur. [1], [7]

Procéder donc comme ce qui va suivre. Pour un nombre impair n donné, choisir aléatoirement un nombre $1 < a < n - 1$. Si $TEST(n)$ retourne « composite » alors n est composite. Si par contre l'algorithme retourne « n pourrait être premier », choisir une autre valeur de a. Si après un nombre assez élever de t tests l'algorithme continue à retourner « n pourrait être premier », assumer que n est premier.

La démonstration de ce qui a été annoncé précédemment (celle qui avance la valeur ¼ comme borne supérieure de la probabilité d'erreur de l'algorithme de Miller-Rabin, pour rappel) s'avère être assez longue. Nous renvoyons ainsi nos lecteurs dans notre bibliographie [7], si besoin est.

3 Conclusion

Cet article nous a permis de se former une solide base sur les fondements mathématiques indispensable à la cryptologie. Pour ne citer que la théorie de l'information, celle des nombres ainsi que les mathématiques discrètes.

Il est quand même assez difficile de couvrir la totalité du domaine. L'on pourrait citer entre autres - les notions de calcul matricielle, la loi des grands nombres, la règle de l'hôpital, l'inégalité de Jensen - notions que nous nous sommes permis d'omettre pour le moment, mais qui ont été utilisées dans nous ouvrages jusqu'ici.

Nous essayerons cependant de trouver une place pour ces notions (en annexe ou autres) dans notre ouvrage final qui consistera en notre thèse doctorale.

4 Bibliographie

[1]. W. Stallings, *« Cryptography and network security, principles and practice, seventh edition »*, Pearson Education Limited, England, 2017 ;

[2]. C. Easttom, *« Modern cryptography, applied mathematics for encryption and information security »*, McGraw-Hill Education, New York, 2016 ;

[3]. K. Bostrom, « *Shannon source coding theorem* », Institut fur physik - Universitat Potsdam, Germany, Apr. 2015 ;

[4]. P. Milne *« Channel Capacity »*, http://www.dip.ee.uct.ac.za/~nicolls/lectures/eee482f/04_chancap.pdf , 2003 ;

[5]. « Entropy (Information Theory) », https://en.wikipedia.org/wiki/Entropy_(information_theory), 2019 ;

[6]. A. Webster and S. Tavares, *« Chapter On The Design Of S-Boxes »*, Electrical Engineering - Queen's University, Kingston Ont. Canada, 1986 ;

[7]. K. Conrad, *« Miller-Rabin Test »*, https://kconrad.math.uconn.edu/blurbs/ugradnumthy/millerrabin.pdf, 2019 ;