Projet : Blockchain appliquée à un processus éléctoral

Nous considérons dans ce projet l'organisation d'un processus électoral par scrutin uninominal majoritaire à deux tours (comme ici en France).

- Partie 1 : Implémentation d'outils de cryptographie.
- Partie 2 : Création d'un système de déclaration sécurisé par chiffrement asymétrique.
- Partie 3 : Manipulation d'une base centralisée de déclarations.
- Partie 4 : Implémentation d'un mécanisme de consensus.
- Partie 5 : Manipulation d'une base décentralisée de déclarations.

Chaque partie est divisée en exercices, qui vont vous permettre de concevoir progressivement le programme final. Il est impératif de travailler régulièrement afin de ne pas prendre de retard et de pouvoir profiter des séances correspondantes pour chaque partie du projet.

Cadre du projet

Dans ce projet, nous considérons la problématique de désignation du vainqueur d'un processus électoral. Dans un processus électoral, chaque participant peut déclarer sa candidature au scrutin et/ou donner sa voix à un candidat déclaré. La tenue d'un processus électoral a, depuis toujours, posé des questions de confiance et de transparence épineuses, dans la mesure où les élections sont généralement organisées par le système exécutif en place, qui est souvent candidat à sa réélection et donc soupçonné d'interférences. De plus, le compte des voix fait appel à des assesseurs, ce qui en fait un travail long et avec peu de garanties de fiabilité, dans la mesure où tout le monde ne peut pas vérifier que le compte a eu lieu dans des conditions honnêtes. Enfin, le caractère anonyme de la désignation par bulletin fait que personne ne peut vérifier a posteriori que sa voix a été comptabilisée chez le bon candidat.

Un autre aspect à considérer est celui de l'ergonomie pour le votant. Plus précisément, un système décentralisé permettrait un vote à distance, et le vote à distance (tout comme le vote par correspondance) a longtemps été envisagé comme un outil pour combattre l'abstention, qui a atteint un record historique lors des des élections régionales et départementales de juin 2021 (66,7 % et allant jusqu'à 87 % chez les jeunes de moins de 25 ans). Un projet de loi a par ailleurs été déposé à l'assemblée nationale le 21 septembre 2021, qui permettrait le vote par correspondance. Ce projet de loi serait un premier pas vers le vote électronique, avec comme argument principal le fait que le vote postal a été instauré en Outre-Rhin en 1957, justifiant possiblement l'écart entre le taux d'abstention en France et celui de l'Allemagne (qui ne dépasse pas les 33%).

L'objectif de ce projet est donc de proposer une piste de réflexion sur les protocoles et sur les structures de données à mettre en place pour permettre d'implémenter efficacement le processus de désignation du vainqueur de l'élection, tout en garantissant l'intégrité, la sécurité et la transparence de l'élection.

Développement d'outils cryptographiques

Dans cette partie, nous allons développer des fonctions permettant de chiffrer un message de façon asymétrique. La cryptographie asymétrique est une cryptographie qui fait intervenir deux clés :

- Une clé publique que l'on transmet à l'envoyeur et qui lui permet de chiffrer son message.
- Une clé secrète (ou privée) qui permet de déchiffrer les messages à la réception.

Par exemple, Bob souhaite envoyer un message à Alice. Pour cela, il utilise la clé publique d'Alice pour chiffrer le message avant de lui envoyer. Une fois reçu, Alice peut utiliser sa clé secrète pour déchiffrer le message. Ces clés peuvent aussi servir à signer des messages (pour vérifier la provenance). Plus précisément, Alice peut utiliser sa clé secrète pour signer des messages qu'elle envoie, et sa clé publique permet aux destinataires de vérifier la signature.

L'algorithme de cryptographie asymétrique que nous allons implémenté est le protocole RSA, très utilisé actuellement sur internet pour transmettre des données confidentielles (notamment dans le cadre du e-commerce). Ce protocole s'appuie sur de (très grands) nombres premiers pour la génération des clés publiques et secrètes. Nous allons donc commencer par traiter le problème de la génération de nombres premiers.

Exercice 1 – Résolution du problème de primalité

Pour générer efficacement des nombres premiers, il faut avoir un moyen d'effectuer rapidement des tests de primalité. Le problème de primalité se définit comme suit : étant donné un entier p impair, p est-il un nombre premier ?

Implémentation par une méthode naïve

Un méthode naïve consiste à énumérer tous les entiers entre 3 et $p-1$, et conclure que p est premier si et seulement si aucun de ces entiers ne divise p.

Q 1.1 Implémentez la fonction `int is_prime_naive(long p)` qui, étant donné un entier impair p, renvoie 1 si p est premier et 0 sinon. Quelle est sa complexité en fonction de p ?

Q 1.2 Quel est le plus grand nombre premier que vous arrivez à tester en moins de 2 secondes avec cette fonction ?

Pour parvenir à générer de très grands nombres premiers, nécessaires au bon fonctionnement du protocole RSA en pratique, nous allons implémenter un test de primalité plus efficace : le test de primalité de Miller-Rabin. Ce test probabiliste nécessite de pouvoir calculer efficacement une exponentiation modulaire, c'est-à-dire la valeur $a^n \bmod m$, étant donnés trois entiers a, n et m. C'est l'objet des questions suivantes.

Exponentiation modulaire rapide

Pour calculer $a^m \bmod n$, sans passer par le calcul de la valeur a^m (qui peut être très grande), une méthode naïve consiste à itérer les étapes suivantes : on multiplie la valeur courante par a puis on applique le modulo n sur le résultat, avant de passer à l'itération suivante. Il s'agit de répéter ces opérations m fois.

Q 1.3 Implémenter la fonction `long modpow_naive(long a, long m, long n)` qui prend en entrée trois entiers a, m et n, et qui retourne la valeur a^m mod n par la méthode naïve. Quelle est sa complexité ?

Q 1.4 Au lieu de multiplier par a à chaque itération, on peut réaliser des élévations au carré (directement suivies de modulo) pour obtenir un algorithme de complexité logarithmique (c-à-d en $O(log_2(m))$). Donnez le code d'une fonction `int modpow(long a, long m, long n)` réalisant cette succession d'élévation au carré.

Q 1.5 Comparez les performances des deux méthodes d'exponentiation modulaire en traçant des courbes de temps en fonction de m. Qu'observez-vous ?

Pour implémenter le test de Miller-Rabin, nous utiliserons la fonction `modpow` ci-après.

Implémentation du test de Miller-Rabin

Le test de primalité de Miller-Rabin est un algorithme randomisé qui utilise la propriété suivante. Soit p un nombre impair quelconque. Soient s et d deux entiers tels que $p = 2^b d + 1$. Soit a un entier strictement inférieur à p. On dit que a est un témoin de Miller pour p si :

- a^d mod $p \neq 1$,
- et $a^{2^r d}$ mod $p \neq -1$ pour tout $r \in \{0, 1, \ldots, b-1\}$.

Si a est un témoin de Miller pour p, alors il est possible de prouver que p n'est pas premier. Malheureusement, dans le cas contraire, on ne peut pas dire que p est premier. Par contre, si on répète ce test suffisamment de fois, pour des valeurs de a tirées au hasard entre 1 et $p-1$, et qu'aucune de ces valeurs générées ne correspond à un témoin de Miller pour p, alors on peut dire que p est très probablement premier.

Ci-dessous, vous trouverez le code des fonctions :

- `int witness(long a, long b, long d, long p)` qui teste si a est un témoin de Miller pour p, pour un entier a donné.
- `long rand_long(long low, long up)` qui retourne un entier `long` généré aléatoirement entre `low` et `up` inclus.
- `int is_prime_miller(long p, int k)` qui réalise le test de Miller-Rabin en générant k valeurs de a au hasard, et en testant si chaque valeur de a est un témoin de Miller pour p. La fonction retourne 0 dès qu'un témoin de Miller est trouvé (p n'est pas premier), et retourne 1 si aucun témoin de Miller n'a été trouvé (p est probablement premier).

```
1
2  int witness(long a, long b, long d, long p) {
3      long x = modpow(a,d,p);
4      if(x == 1){
5          return 0;
6      }
7      for(long i = 0; i< b; i++){
8          if(x == p-1){
9              return 0;
10         }
11         x = modpow(x,2,p);
12     }
13     return 1;
14 }
15
```

```
16 long rand_long(long low, long up){
17     return rand() % (up - low +1)+low;
18 }
19
20 int is_prime_miller(long p, int k) {
21     if (p == 2) {
22         return 1;
23     }
24     if (!(p & 1) || p <= 1) { //on verifie que p est impair et different de 1
25         return 0;
26     }
27     //on determine b et d :
28     long b = 0;
29     long d = p - 1;
30     while (!(d & 1)){ //tant que d n'est pas impair
31         d = d/2;
32         b=b+1;
33     }
34     // On genere k valeurs pour a, et on teste si c'est un temoin :
35     long a;
36     int i;
37     for(i = 0; i < k; i++){
38         a = randlong(2, p-1);
39         if(witness(a,b,d,p)){
40             return 0;
41         }
42     }
43     return 1;
44 }
```

Q 1.6 Intégrer ces fonctions dans votre programme.

On s'intéresse maintenant à la fiabilité du test de Miller-Rabin, autrement dit à sa probabilité d'erreur. L'algorithme fait une erreur quand il déclare qu'un entier p est premier alors qu'il ne l'est pas. Cela se produit quand, pour un entier p non premier, l'algorithme ne trouve pas de témoin de Miller pour p parmi les k valeurs de a générées.

Q 1.7 En utilisant le fait que, pour tout entier p non premier quelconque, au moins $\frac{3}{4}$ des valeurs entre 2 et $p-1$ sont des témoins de Miller pour p, donner une borne supérieure sur la probabilité d'erreur de l'algorithme ?

Comme la probabilité d'erreur de cet algorithme devient rapidement très faible quand k augmente, et que sa complexité pire-cas est en $O(k(log_2(p))^3)$, alors il est plus intéressant d'utiliser cet algorithme pour effectuer des tests de primalité que la méthode naïve implémentée au tout début de ce projet. On utilisera donc le test de primalité de Miller-Rabin dans la suite du projet.

Génération de nombres premiers

On souhaite à présent utiliser le test de Miller-Rabin pour générer des nombres premiers de grande taille, où la taille d'un entier est donnée par son nombre de bits. L'idée est de générer aléatoirement des entiers de la bonne taille (avec la fonction `rand_long` définie précédemment), jusqu'à en trouver un qui réussit le test de Miller-Rabin. Le théorème des nombres premiers assure que l'on trouve par cette méthode un nombre premier au bout d'un nombre d'essais raisonnable.

Q 1.8 Écrire une fonction `long random_prime_number(int low_size, int up_size, int k)` qui étant donnés :

— deux entiers `low_size` et `up_size` représentant respectivement la taille minimale et maximale du nombre premier à générer,
— et un entier k représentant le nombre de tests de Miller à réaliser,
retourne un nombre premier de taille comprise entre `low_size` et `up_size`.

Rappel utile : 2^t est le plus petit entier à t bits, tandis que $2^{t+1} - 1$ est le plus grand. En langage C, un entier `long` correspond à 32 bits.

Exercice 2 – Implémentation du protocole RSA

Le chiffrement RSA, nommé ainsi par les initiales de ses trois inventeurs (Rivest, Shamir et Adleman), est un algorithme de cryptographie asymétrique qui a été décrit en 1977 et breveté en 1983.

Génération d'une clé publique et d'une clé secrète

Pour pouvoir envoyer des données confidentielles avec le protocole RSA, il faut tout d'abord générer deux clés : une clé publique permettant de chiffrer des messages et une clé secrète pour pouvoir les déchiffrer. Pour que les échanges soient sécurisés, le couple (clé secrète, clé publique) doit être engendré de manière à ce qu'il soit calculatoirement impossible de retrouver la clé secrète à partir de la clé publique. Le fonctionnement du protocole RSA est fondé sur la difficulté de factoriser de grands entiers. Plus précisément, pour générer un couple (clé secrète, clé publique), le protocole RSA a besoin de deux (grands) nombres premiers p et q distincts (générés aléatoirement), et réalise les opérations suivantes :

1. Calculer $n = p \times q$ et $t = (p-1) \times (q-1)$.
2. Générer aléatoirement des entiers s inférieur à t jusqu'à en trouver un tel que $\text{PGCD}(s,t) = 1$.
3. Déterminer u tel que $s \times u \bmod t = 1$.

Le couple $pkey = (s,n)$ constitue alors la clé publique, tandis que le couple $skey = (u,n)$ forme la clé secrète. Par définition, u est l'inverse de s modulo t (c'est ce qui permettra le déchiffrement).

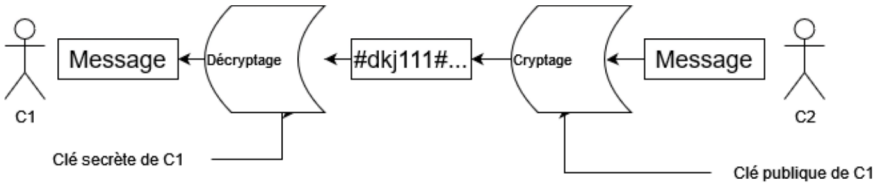
Pour déterminer rapidement la valeur $\text{PGCD}(s,t)$ et l'entier u vérifiant $s \times u \bmod t = 1$, on peut utiliser l'algorithme d'Euclide étendu. En effet, étant deux nombres entiers s et t, cet algorithme calcule la valeur $\text{PGCD}(s,t)$ et détermine les entiers u et v vérifiant l'équation de Bezout : $s \times u + t \times v = \text{PGCD}(s,t)$ En particulier, quand $\text{PGCD}(s,t) = 1$, on a bien $s \times u \bmod t = 1$. Un version récursive de l'algorithme d'Euclide étendu vous est donné ci-dessous :

```
1  long extended_gcd(long s, long t, long *u, long *v){
2      if (t == 0){
3          *u = 1;
4          *v = 0;
5          return s;
6      }
7      long uPrim, vPrim;
8      long gcd = extended_gcd(t, s % t, &uPrim, &vPrim);
9      *u = vPrim;
10     *v = uPrim-(s/t)*vPrim;
11     return gcd;
12 }
```

Q 2.1 Implémentez la fonction `void generate_key_values(long p, long q, long* n, long *s, long *u)` qui permet de générer la clé publique $pkey = (s, n)$ et la clé secrète $skey = (u, n)$, à partir des nombres premiers p et q, en suivant le protocole RSA.

Chiffrement et déchiffrement de message

On s'intéresse maintenant à l'envoi de message. Supposons que la personne C2 souhaite envoyer un message à la personne C1 en utilisant le protocole RSA. Dans ce cas, la personne C2 utilise la clé publique de la personne C1 pour chiffrer le message avant son envoi. À sa réception, la personne C1 déchiffre le message à l'aide de sa clé secrète.



Soient $pkey = (s, n)$ et $skey = (u, n)$ la clé publique et la clé secrète du destinataire, et m le message à lui envoyer représenté par un entier (inférieur à n).

- **Chiffrement :** on chiffre le message m en calculant $c = m^s \bmod n$ (c est la représentation chiffrée de m).
- **Déchiffrement :** on déchiffre c pour retrouver m en calculant $m = c^u \bmod n$.

Q 2.2 Implémentez une fonction `long* encrypt(char* chaine, long s, long n)` qui chiffre la chaîne de caractères `chaine` avec la clé publique $pkey = (s, n)$. Pour cela, la fonction convertit chaque caractère en un entier de type `int` (sauf le caractère spécial '\0'), et retourne le tableau de `long` obtenu en chiffrant ces entiers.

Rappel : il faut utiliser la fonction `modpow` pour réaliser une exponentiation modulaire efficace.

Q 2.3 Implémentez une fonction `char* decrypt(long* crypted, int size, long u, long n)` qui déchiffre un message à l'aide de la clé secrète $skey = (u, n)$, en connaissant la taille du tableau d'entiers. Cette fonction renvoie la chaîne de caractères obtenue, sans oublier le caractère spécial '\0' à la fin.

Fonction de test

Pour vérifier le bon fonctionnement de votre programme, reproduisez le programme principal suivant et compilez le en utilisant les fonctions que vous avez implémentées dans les questions précédentes.

```
1  void print_long_vector(long *result, int size){
2     printf("Vector: [");
3     for (int i=0; i<size; i++){
4        printf("%ld \t", result[i]);
5     }
6     printf("]\n");
7  }
8
9
10 int main()
11 {
12    //Generation de cle :
```

```
13      long p = random_prime_number(15,16, 5000);
14      long q = random_prime_number(15,16, 5000);
15      long n, s, u;
16      generate_keys(p,q,&n,&s,&u);
17      printf("cle publique = (%ld, %ld) \n",s,n);
18      printf("cle privee = (%ld, %ld) \n",u,n);
19
20      char message[1000] = "Hello";
21      int len = strlen(message);
22
23      //Chiffrement:
24      long* crypted = encrypt(message, s, n);
25
26      printf("Initial message: %s \n", message);
27      printf("Encoded representation : \n");
28      print_long_vector(crypted, len);
29
30      //Dechiffrement
31      char* decoded = decrypt(crypted, len, u, n);
32      printf("Decoded: %s\n", decoded);
33
34      return 0;
35  }
```