RAPPORT

Sujet : CRYPI
Groupe3

Groupe N°3

Lucas SIAUVE
Clément BRUN
Elsa FRANCOIS
Matthieu TIRLOY

Table des matières

1 Introduction

Ce rapport présente ce que nous avons réalisé pour la conception et l'implémentation de notre un mécanisme de vérification de preuve à divulgation nulle de connaissance (ZKP). C'est un mécanisme qui permet de faire de la vérification d'identité, d'âge, et autres caractéristique personelles, tout en ne dévoilant aucune connaissance à l'acteur qui souhaite les vérifier. Dans notre société où la divulgation de données personelles et sensibles est devenue un point d'intérêt pour la majorité des utilisateurs, nous proposons une implémentation de cette méthode où les attributs d'identité, engagés par une autorité de confiance via un hachage cryptographique, peuvent être prouvés sans révéler d'informations suplémentaires. Ce projet nous permet finalement d'explorer les fondements de la cryptographie, de réaliser une mise en œuvre pratique à l'aide d'une bibliothèque ZKP, et d'analyser ensuite les forces et limites de notre système.

2 Le modèle d'engagement d'identité

2.1 Définition

Le modèle d'engagement d'identité consiste à représenter de manière sécurisée et confidentielle les attributs personnels d'un individu sous forme d'un engagement cryptographique. Cet engagement sert de preuve immuable des données d'identité sans nécessiter leur divulgation directe.

Une autorité de confiance (le plus souvent une organisation apparentée au gouvernement) collecte les attributs d'identité d'un utilisateur. Dans notre cas, les attributs sont définis : le nom, le prénom, la date de naissance, la catégorie du permis de conduire, la date d'expiration et un nonce qui est aléatoire. Ces données sont ensuite concaténées et soumises à un algorithme de hachage selon le format prédéfini suivant :

$$H(\texttt{name} \parallel \texttt{surname} \parallel \texttt{date\_of\_birth} \parallel \texttt{expiration\_date} \parallel \texttt{nonce})$$

Ce modèle d'engagement présente plusieurs propriétés intéressantes et essentielles :

- **Confidentialité :** La fonction de hachage étant à sens unique, il est impossible de retrouver les données originales à partir de l'engagement seul, protégeant ainsi directement les données de l'utilisateur.
- **Intégrité :** Toute modification des attributs d'identité entraînerait une modification du hash final, garantissant ainsi une intégrité des données fournies.
- **Non reproductible :** L'ajout d'un nonce comme une valeur aléatoire non divulguée évite que deux utilisateurs avec des données similaires produisent le même engagement, ou même qu'un attaquant parvienne à retrouver les données de l'utilisateur en brute forçant le hash en essayant plusieurs combinaisons de données en entrée.

2.2 Réflexions sur les impacts la publication publique de l'engagement

Quels sont les avantages pour la protection de la vie privée d'une telle publication ?

- **Protection des données personnelles :** Le hachage agit comme une preuve numérique non réversible et non altérable (intégrité conservée). Il permet de prouver que des données existent, sont vérifiables par une preuve, tout en n'étant pas dévoilées au public.
- **Vérification décentralisée :** Toute entité disposant du hachage peut vérifier une preuve sans avoir accès aux données originales. La confiance qu'elle accorde à la preuve vient directement de l'autorité de vérification, qui est la seule à recevoir les données sensibles dans le processus de vérification.
- **Transparence et auditabilité :** L'engagement étant public, il est vérifiable par des tiers sans compromission de la vie privée. Cela permet une transparence quasi totale dans la gestion des identités, et surtout dans le respect des données personelles et sensibles.

Quels en sont les risques ?

- **Corrélation entre engagements :** Sans l'ajout d'un nonce unique, aléatoire et suffisament grand, deux engagements issus de données identiques généreraient le même hash. Cela rendrait possibles des corrélations entre utilisateurs, dévoilant ainsi à certains utilisateurs les données personelles d'autres utilisateurs.
- **Attaques par dictionnaire ou force brute :** Si certaines données (comme le nom ou la date de naissance) sont prévisibles voire connues, un attaquant pourrait tenter de reconstituer l'engagement en testant toutes ou une série de combinaisons possibles. À nouveau, sans l'utilisation d'un hash suffisamment robuste, cela compromettrait dangereusement les engagements publiés, qui pourraient être "brisés" par calculs de hash en série par exemple.
- **Fuite d'information par méta-données :** Cette hypothèse, moins probable mais tout de même envisageable, suppose que des informations sur les utilisateurs pourraient être divulgués lors des processsus et manipulation des données de l'autorité de confiance. Elle deviendrait le maillon faible de la chaîne et la cible privilégiée des attaquants. Le moment de la publication, la structure des engagements ou l'ordre d'insertion dans la base publique sont autant d'éléments qui pourraient potentiellement révéler des indices sur les utilisateurs.

Sous quelles conditions ce modèle peut-il être considéré comme sûr ?

Pour que ce modèle soit considéré comme sécurisé, plusieurs conditions doivent être réunies :

- **Fonction de hachage robuste :** La confidentialité des données repose premièrement sur laLa fonction de hachage utilisée doit être résistante aux collisions, aux pré-images et aux secondes pré-images (exemples : SHA-256, Blake2).
- **Utilisation d'un nonce cryptographiquement sécurisé :** Le *nonce* doit être aléatoire, de taille suffisante, et différent pour chaque engagement afin de garantir l'unicité et empêcher les attaques par tables arc-en-ciel.
- **Canal sécurisé pour la transmission des preuves privées :** Bien que l'engagement soit public, les éléments utilisés dans les preuves doivent rester strictement privés et confidentiels côté utilisateur.
- **Pas de fuite par les entrées publiques :** Les données publiques utilisées pour la vérification (ex. nom, prénom) doivent être choisies de manière à ne pas permettre de reconstituer les attributs privés.

3 Critères et processus de sélection de la bibliothèque ZKP

Le choix de la bibliothèque ZKP s'est porté sur CRICOM, en fonction des critères suivants :

- **Simplicité d'utilisation** : interface claire, bonne documentation, adaptée à un contexte pédagogique.
- **Fonctionnalités adaptées** : prise en charge des opérations nécessaires (comparaison d'âge, égalité de catégories, recalcul de hash).
- **Compatibilité avec Poseidon** : fonction de hachage efficace et optimisée pour les circuits ZKP.
- **Pipeline intégré** : compilation, génération et vérification de preuve simples, sans configuration complexe.
- **Preuves efficaces** : protocoles comme Groth16 pour produire des preuves courtes et rapides à vérifier.
- **Facilité d'expérimentation** : permet un prototypage rapide et une bonne intégration dans notre architecture.

4 Résultats des expériences menées

Nous avons rencontré un souci lié à la licence et au permis à cause de la fusion des deux circuits. Sinon, pour le résultat indiquant si la personne est majeure, nous avons ceci :

- 1 si elle est majeure,
- 0 si elle est mineure,
- 2 si le hash ne correspond pas.

5 Défis rencontrés au cours du projet

Le passage où nous vérifions si le hash est correct a été compliqué. Nous avions d'abord commencé par réaliser des tableaux pour les chaînes de caractères, que nous avons ensuite transformés en un hachage de nombres avec Poseidon. Nous avons essayé de fusionner deux circuits différents : un pour l'âge, l'autre pour le permis. Nous n'avons pas réussi, mais nous avons laissé le circuit circom à l'intérieur du projet. Nous avons aussi rencontré des problèmes pour fusionner l'interface web avec le circuit.

6 Étapes futures recommandées pour améliorer le projet

Fusionner les différents circuits pour faciliter la vérification de ce que nous voulons a été un défi. Nous avons fait des circuits séparés, mais peut-être serait-il mieux d'en créer un seul, plus complet, qui contienne tout, comme pour une pièce d'identité.

7 Conclusion

La garantie de confidentialité offerte par le système ZKP repose sur la capacité à prouver la validité d'un attribut (comme être majeur ou avoir un permis valide) sans révéler les données sensibles sous-jacentes. Tant que la fonction cryptographique Poseidon reste sécurisée, les données originales restent protégées, et l'utilisateur ne peut pas être identifié via les engagements publics.

Cependant, les engagements basés sur un hash présentent des limites, notamment leur difficulté à gérer des intervalles complexes et leur dépendance à la qualité du nonce. De plus, la confiance repose sur un émetteur unique, ce qui peut être un point de faiblesse en cas de compromission.

Techniquement, le système est adapté à une échelle nationale grâce à la compacité des engagements et des preuves, mais son déploiement nécessite une infrastructure robuste pour la gestion des clés, la mise à jour des circuits, et un service d'API fiable.

D'autres approches cryptographiques, comme les certificats d'attributs signés, les identifiants sélectifs ou les signatures aveugles, peuvent offrir des solutions complémentaires ou alternatives mieux adaptées selon les besoins.

En résumé, le système ZKP présente une forte protection de la vie privée et une faisabilité technique prometteuse, mais son succès dépendra d'une mise en œuvre rigoureuse et de la prise en compte de ses limites intrinsèques.