Rapport de Stage

Jewin CHENG LDD3 Magistère d'Informatique
Sous la direction de Sylvain Conchon

Mai - Juillet 2024

Table des matières

1 Introduction

Ce stage, réalisé au sein du Laboratoire des Méthodes Formelles (LMF) de mai à juillet, avait pour objectif d'initier l'étudiant aux concepts de la blockchain, avec un focus particulier sur Monero.

1.1 Blockchain

La blockchain est une technologie de registre distribué où les transactions entre utilisateurs sont enregistrées. Chaque utilisateur doit disposer d'un montant nécessaire, représenté par d'autres transactions reçues, pour créer une nouvelle transaction et envoyer une certaine quantité de cryptomonnaie à un autre utilisateur. Les transactions sont regroupées en blocs, chaque bloc étant lié au précédent par un hash cryptographique, formant ainsi une chaîne continue depuis le premier bloc, appelé "genesis". N'importe quel utilisateur à également la possibilité de voir l'ensemble des transactions passées, leur montant et les acteurs impliqués dans ces transactions.

Ce système est décentralisé et vérifié par un réseau de nœuds (ordinateurs) qui partagent et valident les transactions. Une fois qu'un certain nombre de blocs a été ajouté à la chaîne, la blockchain devient pratiquement immuable, ce qui garantit l'intégrité et la sécurité des données. Pour ajouter de nouveaux blocs, le réseau doit parvenir à un consensus, un processus qui varie en fonction des mécanismes de consensus utilisés, comme la Preuve de Travail (PoW) ou la Preuve d'Enjeu (PoS).

Les nouveaux blocs sont "minés", c'est-à-dire que leur hash est calculé par un ensemble d'ordinateurs qui résolvent des problèmes mathématiques complexes. Le hash du bloc doit respecter certaines règles pour être valide. Ce processus de minage est également le moyen par lequel de nouvelles unités de cryptomonnaie sont créées et introduites dans le système.

1.2 Monero

Dans le cadre de ce stage, l'attention a été portée sur les protocoles garantissant la sécurité et l'anonymat des utilisateurs de Monero, une cryptomonnaie axée sur la confidentialité lancée en 2014. Les principaux protocoles étudiés incluent la signature en anneau, les adresses furtives, les clés images et Ring Confidential Transactions (RingCT), qui combinent les trois pour masquer les détails des transactions. Ce sont ces protocoles qui rendent Monero différent des autres blockchains comme Bitcoin.

1.3 Objectif

L'objectif principal était de comprendre ces mécanismes de sécurité et d'anonymat, et d'explorer la possibilité de tracer les utilisateurs de Monero. L'un des fils conducteurs de ce stage était un papier de recherche de 2017 qui présentait deux faiblesses sur l'anonymat des utilisateurs de Monero qui seront expliqués dans la prochaine partie. Bien que des résultats concrets en termes de traçabilité n'aient pas été atteints, plusieurs outils et ressources ont été développés, notamment un outil de visualisation de Monero sous forme de graphe basé sur des intervalles de hauteur de blocs, ainsi que des slides de présentation et des statistiques sur les transactions.

Ce rapport se concentre sur les transactions entre utilisateurs et les protocoles cryptographiques assurant leur sécurité et anonymat dans Monero.

2 État de l'art

Le début du stage consistait à lire un papier de recherche intitulé "An Empirical Analysis of Linkability in the Monro Blockchain" à propos de la traçabilité des utilisateurs de Monero. Dans ce papier datant de 2017, il était décrit que Monero avait des failles en ce qui concerne l'anonymat de ses utilisateurs.

2.1 Une transaction dans la blockchain

Pour qu'un utilisateur puisse envoyer de l'argent à un destinataire, l'utilisateur doit créer une transaction en utilisant une qu'il a reçue qui n'a pas encore été dépensé (UTXO) dans une autre transaction. Cette transaction était ensuite soumis dans le réseau qui le vérifie pour être finalement intégré dans un bloc et ainsi acté cet envoie de cryptomonnaie. Comme le registre de blocs est distribué, n'importe qui peut les voir et donc potentiellement savoir quel utilisateur a reçue tant de cryptomonnaie et envoyé tant.

2.2 Méthode d'anonymisation : le mixage

Des méthodes ont été développés dans le but d'anonymiser la connexion entre l'envoyeur et le destinataire de cryptomonnaie, vaguement utilisé par les criminels. L'une de ces méthodes est le "mixng" qui consiste à réunir un grand nombre d'utilisateur voulant envoyer un certain montant à leur destinataire. Cela a été pensé notamment pour Bitcoin où des plateformes proposaient ce genre de services. N'importe qui voit lors d'un mixage que, plein de transactions ayant le même montant avait comme destination la plateforme et la plateforme envoie à tous les destinataires le même montant. De ce fait personne ne savait qui avait envoyé exactement à qui, seulement qu'un utilisateur à envoyé un montant précis à une autre personne.

2.3 Le mixage dans Monero

Monero propose cette fonctionnalité sans passer par un tiers comme une plateforme mais directement intégré dans la création d'une transaction avec leur protocole, la signature en anneau. Le fonctionnement est légèrement différent, une seule transaction est utilisé et le mixage est constitué de transactions issues de n'importe quel utilisateur avec le même montant que celle réellement utilisé.

2.4 Résultat du papier de recherche

Cependant, le papier de recherche montre qu'il était possible de déterminer les transactions réellement utilisé dans le mixage. Deux failles ont été montrés, la première étant que le mixage n'était pas obligatoire et donc certaines transactions n'incluait pas de mixage. On pouvait donc être certains que dans ce type de transaction, celle

utilisé était réellement utilisé. Sauf que ces transactions qui ne sont plus utilisable (non UTXO) étaient inclus dans d'autres mixages, il était alors possible de déduire par élimination quelle transaction était réellement utilisé dans ce mixage. Une autre faiblesse étant que chaque transaction mixeur doit être du même montant que celle réellement utilisé, réduisant ainsi le nombre de mixeurs disponible. La deuxième faille était que les transactions utilisés pour le mixage étaient souvent les plus récentes, la transaction réellement utilisé était dans 80% des cas la plus récente. Ces transactions dont on pouvait déduire celle dépensé constituait 62% des transactions totales.

2.5 État actuel de Monero

En réponse à ces problèmes, Monero a effectué un bon nombre de mise à jour incluant RingCT et l'obligation d'inclure un minimum de mixing en 2018 (à ce jour il est de 15, il est prévu de monter jusqu'à 100 mixeurs possibles d'après leur roadmap). Les transactions qui implémentent le protocole RingCT ne peuvent utiliser que des transactions qui l'implémentent également ce qui fait que les transactions déductibles et n'ayant pas de mixage n'impactent plus les transactions modernes. Aujourd'hui les heuristiques de ce papier ne sont plus forcément pertinentes pour répondre au problème de traçabilité puisque les transactions n'ont plus les failles pointés par ces heuristiques.

3 Problèmes traités et solutions proposées

3.1 Problèmes traités

Le problème traité dans ce stage était d'abord de comprendre les différents protocoles de sécurité de Monero (signature en anneau, adresses furtives, engagements de Pedersen et RingCT) et leur impact sur les heuristiques définies dans un papier de recherche. Par la suite, il s'agissait d'explorer ce qu'il était possible de faire en prenant en compte les nouveaux protocoles.

Divers documents et discussions sur des forums ont été consultés pour comprendre le fonctionnement des différents protocoles. Après quelques échanges avec l'encadrant, nous sommes finalement parvenus à comprendre leur fonctionnement. Les détails des protocoles sont décrits dans les slides de présentation.

Pour résumer, la signature en anneau permet de masquer l'identité de l'utilisateur en mélangeant la transaction utilisée avec d'autres transactions appelées mixeurs. Différentes versions sont détaillées dans le livre "Zero to Monero", la plus récente étant CLSAG. Les adresses furtives permettent de masquer le destinataire en créant une nouvelle adresse à partir de la clé publique du destinataire. Un secret partagé entre l'envoyeur et le receveur permet le bon fonctionnement de l'échange. Les engagements de Pedersen masquent les montants des entrées et sorties grâce à ce même secret. Tous ces protocoles réunis dans RingCT rendent la traçabilité des utilisateurs complexe.

3.2 Solutions proposées

En réponse à cette problématique de traçabilité, un outil pour visualiser le mixing dans Monero sous forme de graphes a été développé, ainsi qu'une petite interface encore améliorable pour visualiser des statistiques sur les transactions.

Étant donné le nombre conséquent de mixeurs par transaction, il n'était pas envisageable de créer un graphe prenant en compte toutes les entrées et sorties. Le choix a été fait de ne garder que les nœuds présents dans plusieurs transactions, présentés comme dépensés par le dataset blackball (qui recense uniquement les adresses post-RingCT de 2017 à 2019). Plus l'intervalle de blocs était grand, plus les mixeurs se connectaient à un grand nombre de transactions, ce qui faisait que le graphe conservait de plus en plus de nœuds et d'arêtes. C'est pourquoi le nombre de transactions connectées est supérieur à deux (ce paramètre peut être ajusté).

Les statistiques, combinées avec des sites en ligne pour explorer en détail une transaction comme ExploreMonero ou LocalMonero (qui cache les transactions blackball), montrent que les entrées blackball dataient d'avant 2020 et provenaient souvent de transactions de minage. Les récentes transactions incluent la plupart des mixeurs de la même année. Il se pourrait que la distribution des entrées mixeurs soit toujours déséquilibrée, comme l'indiquait la deuxième heuristique du papier de recherche. Cependant, le manque d'informations sur la réelle transaction dépensée ne permet pas de vérifier cela.
Le manque de temps et d'espace mémoire étaient également des facteurs limitant pour approfondir cette analyse.

4 Conclusion et perspectives

4.1 Conclusion sur le sujet

La traçabilité de Monero est complexe, ce qui en fait une cryptomonnaie plus sécurisée que d'autres. Cela la rend très utile pour les criminels, bien qu'il soit peut-être possible de les attraper.

Une cryptomonnaie, pour être utile, doit nécessairement interagir avec les monnaies ou objets du monde réel, que ce soit à travers des achats ou la conversion de monnaie, ce qui se fait par le biais de plateformes. En collaborant avec ces plateformes, il devient possible d'obtenir des adresses qui semblent appartenir à des individus malveillants. Bien qu'il soit difficile de remonter jusqu'à l'individu à cause des différents protocoles, on a vu qu'il était possible de trouver des failles, par exemple avec la signature en anneau.

4.2 Perspectives d'améliorations

Il serait intéressant de vérifier la distribution d'âges des mixeurs si la transaction réellement utilisée est connue, comme dans le papier de recherche. Certaines transactions utilisent encore de vieilles transactions datant de 2017 ou 2018, époque où le protocole RingCT a été implémenté, créant un fossé entre les transactions utilisant ce protocole

et celles qui ne l'utilisent pas. Ces vieilles transactions proviennent souvent directement du minage. Il serait intéressant de voir si ces transactions de minage sont mélangées dans d'autres transactions récentes. Il y a encore des transactions issues des protocoles antérieurs à RingCT dans les blocs modernes.

D'autres statistiques sur la quantité de monnaie créée, le poids ou le nombre total d'entrées pourraient être des indicateurs pour rendre une transaction suspecte. Plus il y a d'entrées et plus il y a de monnaie utilisée à destination de quelqu'un, ce qui pourrait être un signe de blanchiment d'argent. Comme les montants sont cachés, il est probable que des transactions suivant une structure commune soient également suspectes et impliquent des sommes importantes.

De toute évidence, augmenter l'intervalle de blocs sur plusieurs années serait d'une grande aide, ne serait-ce que pour vérifier les idées mentionnées plus haut, mais aussi pour en trouver d'autres.

4.3 Conclusion personnelle

Pour conclure de manière plus personnelle, j'aimerais avant tout remercier Monsieur Sylvain Conchon pour l'opportunité qu'il m'a offerte de découvrir un sujet en vogue et de m'avoir introduit au monde de la recherche, les séminaires, l'environnement et les personnes rencontrées. Cela m'a permis de progresser sur plusieurs aspects techniques et de gagner en confiance pour mon parcours futur. J'ai beaucoup apprécié cette expérience, qui fut très enrichissante et plaisante. Si je devais la conseiller à quelqu'un, je n'hésiterais pas une seule seconde.

J'aimerais également remercier mes camarades du magistère (promo L3 2024-2025), qui ont été mes voisins de bureau durant ce stage, qui m'ont aidé sur ce sujet et avec qui j'ai partagé de nombreux bons moments (notamment au babyfoot et au billard), sans oublier Madame Alexandrina Korneva, qui m'a aidé sur ce sujet, notamment en m'aiguillant sur les choix de conception des outils.

Finalement, j'aimerais remercier Monsieur Thibaut Balabonski, qui m'a poussé à soumettre ma candidature au magistère et sans qui je n'aurais pas eu la chance de réaliser cette année aux côtés de personnes aussi talentueuses que bienveillantes.

Bibliographie

Ressources et outils utiles

Canva: Slides de présentation sur les protocoles cryptographiques de Monero
Fichiers pour faire un blackball dataset
Test de Monero sur navigateur
GUI pour créer un wallet dans Monero
Génération de clé
Daemon rpc

Références

Monero roadmap
An Empirical Analysis of Linkability in the Monro Blockchain
Zero to monero (version 2)
Monero research lab, papiers sur les différents protocoles de Monero
Signature en anneau
Exemple d'engagement de Pedersen
Bulletproof+
Exemple sur la création de clé et les courbes elliptiques
Exemple adresses furtives
Doublé dépense avec les clés images
Conversion des keyoffsets

Ressources à explorer

Projet sur la signature en anneau