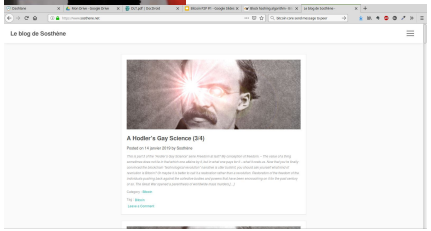
Bitcoin Pair-à-pair

#2 - Confidentialité

Objectifs

1. La confidentialité dans Bitcoin : en quoi suis-je concerné ?
2. Comment peut-on me désanonymiser ?
3. Quels sont les moyens de m’en protéger ?

www.sosthene.net
@Sosthene@bitcoinhackers.org



1. Pourquoi la confidentialité ?

La confidentialité est essentielle pour...

1. La fongibilité de Bitcoin : une bonne monnaie est fongible par définition
2. L’activité normale des acteurs économiques sur le marché libre
3. La sécurité de chacun contre les vols et les attaques diverses
4. et tout simplement, parce que vous avez le droit de choisir quelles informations vous voulez partager (ou non !) et avec qui

Dernier point, la confidentialité n’est en rien incompatible avec les missions de police et de justice, cela est parfaitement compris dans le monde physique et il n’y a pas de raison qu’il en soit autrement dans le numérique.

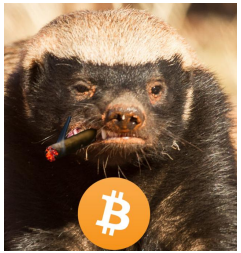
Bitcoin est-il “non-confidentiel” par nature ?

- Il y a quelques années, on considérait souvent que Bitcoin était “anonyme” (d’où le discours sur l’argent des criminels et des terroristes)
- Aujourd’hui, le discours a radicalement changé, et on entend au contraire que Bitcoin est totalement transparent et qu’on peut identifier tout ce qui s’y passe.
- Comme souvent la vérité est beaucoup plus nuancée...

Les dilemmes de Bitcoin

- Pseudonyme vs anonyme
  - Contrairement à d'autres cryptos, dans Bitcoin le payeur, le destinataire et le ou les montants sont publics
  - Cela facilite l'audit par tous de la block chain
- Audit décentralisé vs confidentialité
  - Tout doit être visible par tous pour toujours
  - Chacun doit pouvoir l'utiliser avec une mesure acceptable de confidentialité
- P2P vs Custody
  - Bitcoin est en-soi "pair-à-pair", un pair se définit par son autonomie
  - L'écosystème fournit aujourd'hui principalement des services qui conservent les bitcoins en dépôt
- On-chain vs off-chain
  - Une transaction n'est réellement valide que si elle est enregistrée "on-chain"
  - Une transaction ne peut être 100% confidentielle que si elle n'est jamais enregistrée

Malentendu #1 : “coin” vs “UTXO”



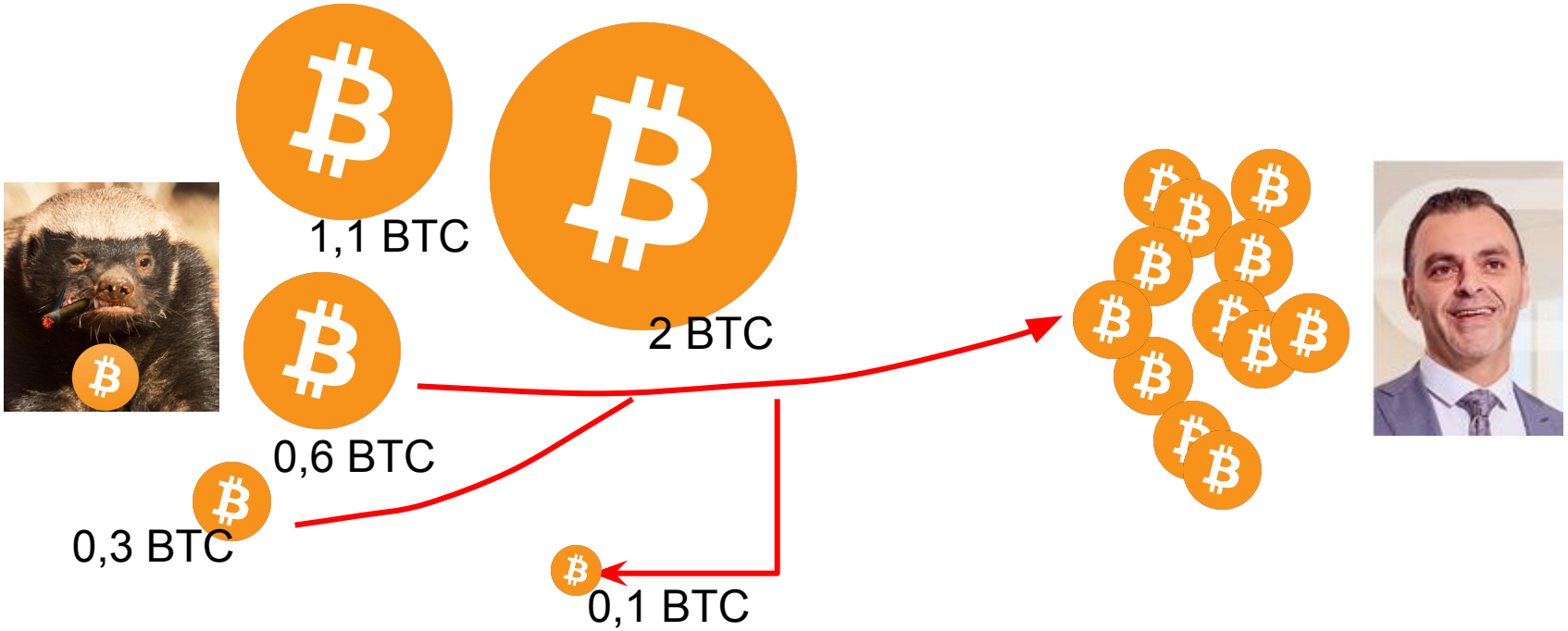
1,1 BTC

2 BTC

0,6 BTC

0,3 BTC

Malentendu #1 : “coin” vs “UTXO”



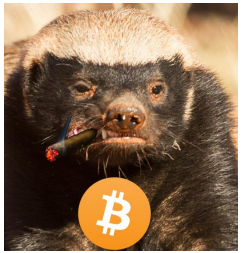
“bitcoin-cli listunspent”


```
        "txid":
"0312dbaf4453063d98269152851579d447f47e12cc1ee0
409361c9c856729506",
        "vout": 0,
         "address":
"mgxnU18MQmxUDmxny5a4pUMhn1mX24Vkp3",
        "scriptPubKey":
"2103a7d88eee6b2c4a782bcd26fca227d056d78883cf6e
de13b64523d67351bdbea4ac",
         "amount": 25.00000000,
         "confirmations": 106,
         "spendable": true,
         "solvable": true,
         "safe": true
```

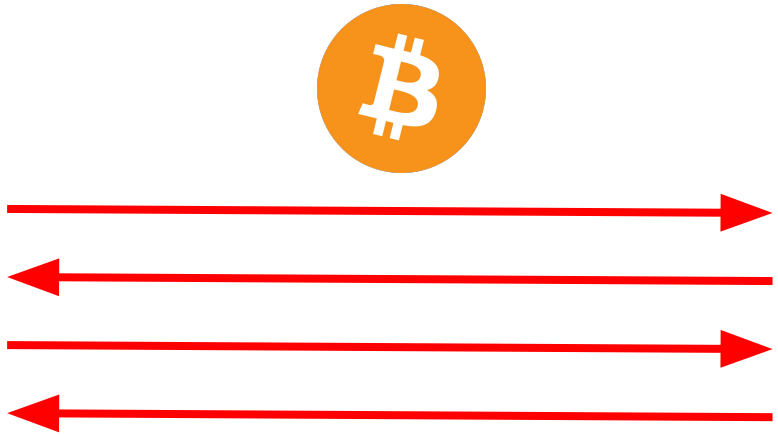
“bitcoin-cli decoderawtransaction”

https://gist.github.com/BobleChinois/6d1541205c596936006cf6864cb2e25e

Malentendu #2 : une “adresse” bitcoin, c’est comme une adresse mail

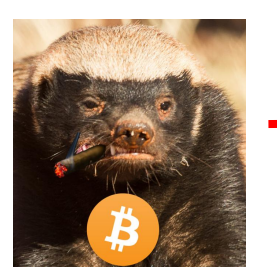


=
1foo...badger
=
4 BTC



=
1foo...scammer
=
2 BTC

Malentendu #2 : une “adresse” bitcoin, c’est comme une adresse mail



=

1foo...badger

=

?

Transaction x

?

1foo...scammer

=

2 BTC de la transaction x

Le mot “adresse” est très mal choisi

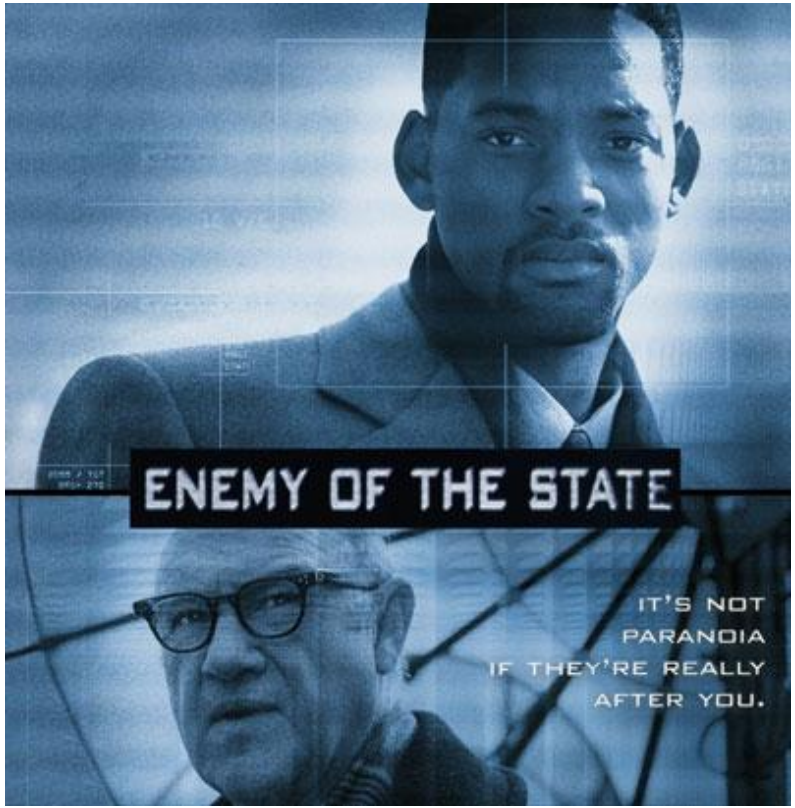
- Une adresse n’est pas réutilisable
- Une adresse n’a pas de balance associée
- Il n’y a pas de notion d’adresse d’expéditeur

Une adresse n’est en fait qu’une représentation standardisée d’une clé publique (ou d’un script) générée pour recevoir un paiement, et devrait plutôt être appelée “facture”.

Les règles des transactions Bitcoin

1. Chaque input est consommé dans son intégralité
2. Le paiement et la “monnaie” font partie de la même transaction
3. Une transaction peut comporter un nombre arbitraire d’inputs et d’outputs **qui n’ont pas de liens déterminés entre eux**
4. Un “portefeuille” ne contient rien qui ressemble même de loin à des pièces, mais les coordonnées des transactions précédentes et des clés privées

2. Quel modèle de menaces ?



Les attaques de désanonymisation

Elles sont de deux types :

1. Certaines exploitent les données de la block chain (“blockchain analysis”)
2. La plupart exploitent toutes les sources informations et ne sont pas spécifique à Bitcoin (trafic internet, KYC, traces laissés sur les réseaux sociaux...)

Toutes ces techniques sont évidemment **cumulables**: une seule d’entre elle ne suffira probablement pas à vous identifier, en revanche **si vous multipliez les traces et les indices, l’accumulation peut être dévastatrice** !

Analyse de block chain

- Il s’agit d’une **heuristique** (c’est-à-dire, tout sauf une science exacte !)
- Elle repose sur un petit nombre **d’hypothèses** qui permettent d’interpréter la structure des transactions sur la block chain
- Elle ne permet pas de rattacher des transactions à une quelconque personne physique, car les données de la block chain sont pseudonymes: **elle vise à constituer des “clusters” (“regroupements”) d’adresses et d’outputs présumés appartenir à la même personne ou entité**.
- Si votre identité est rattachée à un cluster, le résultat peut être catastrophique

Principales hypothèses

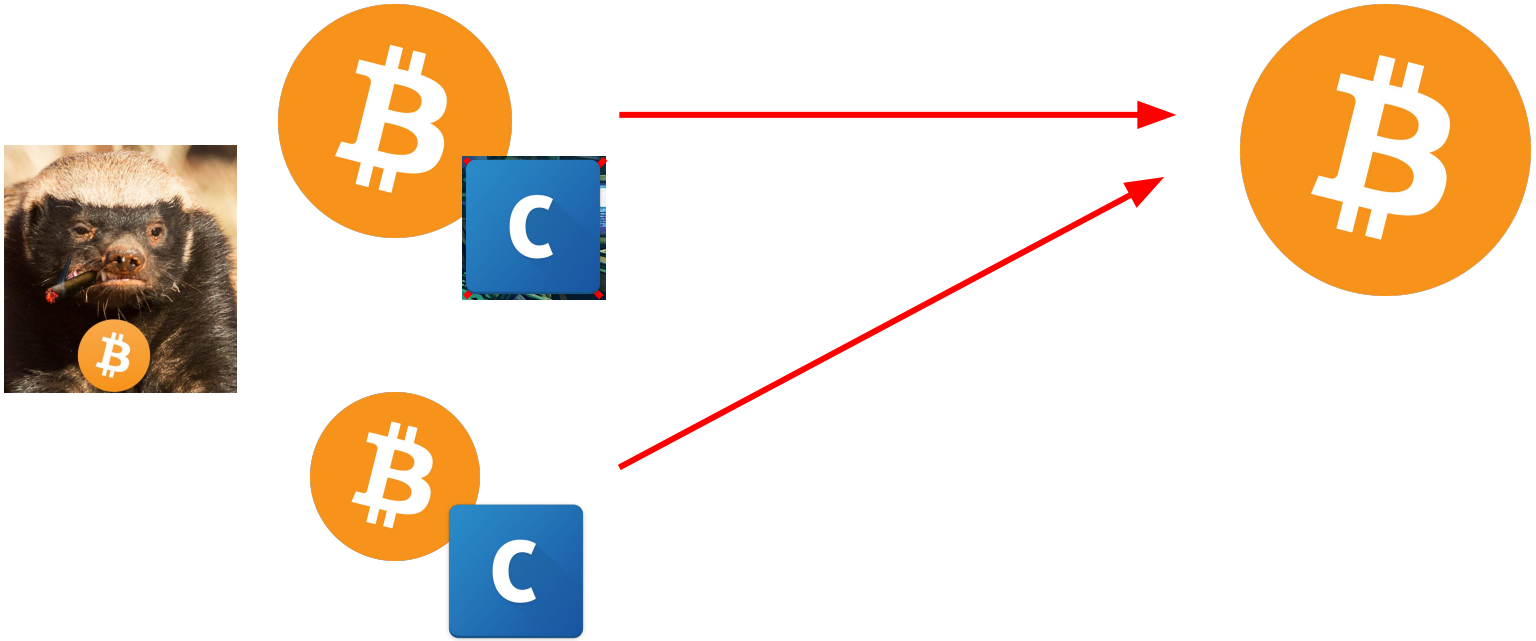
- **Common-input-ownership heuristic** : tous les inputs d'une transaction appartiennent à la même entité.
- **Change address detection** : parmi les outputs, l'un correspond à la "monnaie", et appartient à la même entité que les inputs.
- **Exact payment amount** : Si une transaction consomme l'intégralité d'un ou de plusieurs UTXO sans générer de monnaie, le destinataire du paiement est identique au payeur (paiement à soi-même).

Techniques actives

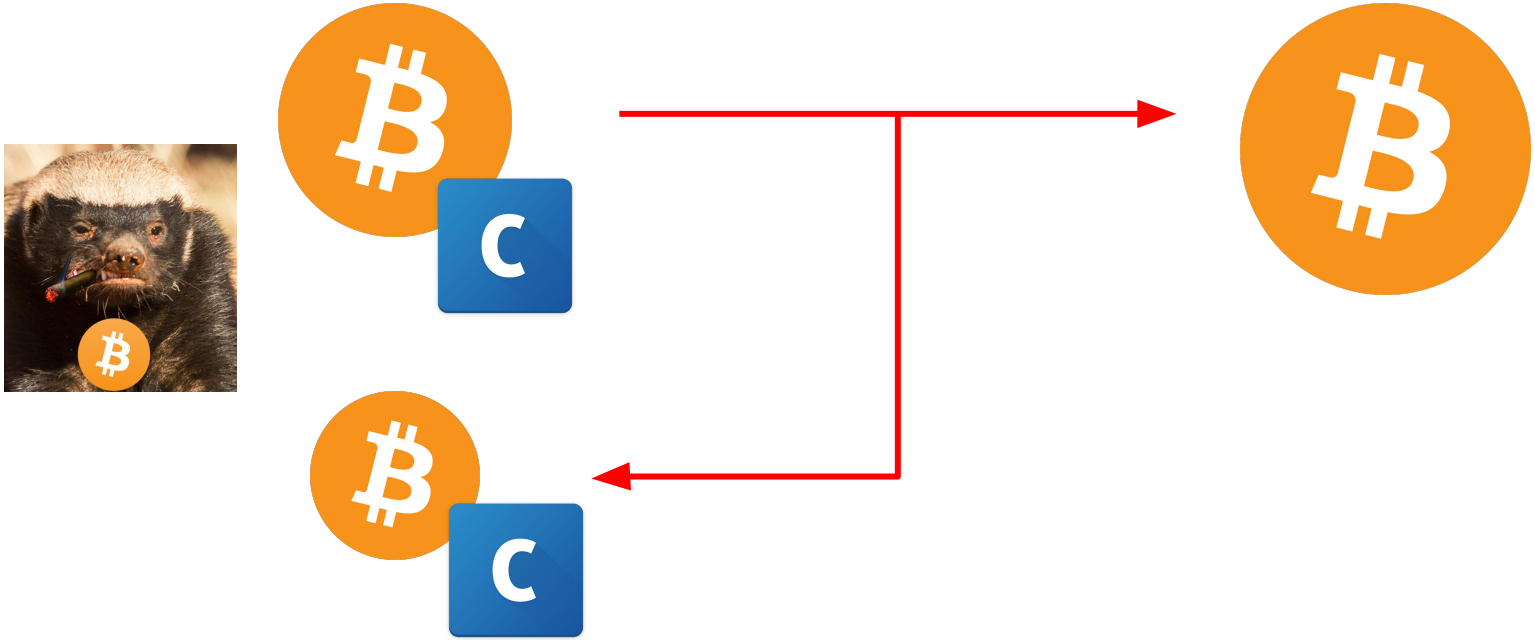
L’attaquant peut ne pas se limiter à observer passivement les transactions et agir pour augmenter l’efficacité de son analyse :

- **Mystery shopper payment** : l’attaquant envoie une petite somme sur une adresse qu’il sait appartenir à sa cible, ce qui lui donne un point de départ ferme pour commencer l’analyse.
- **Forced address reuse** : l’attaquant envoie une petite somme sur des adresses déjà utilisées, dans l’espoir que la cible dépensera cet output en même temps qu’un autre ce qui permettra de lui associer d’autres adresses.

Common-input-ownership

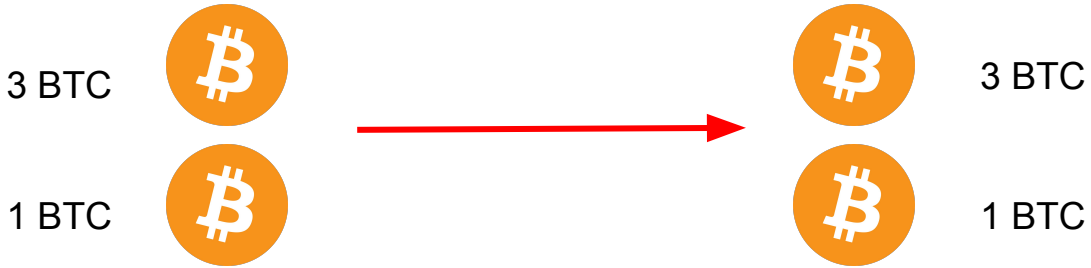


Change-address detection



Quel degré de certitude de l’analyse ?

- Prenons une transaction très simple :



- Combien de façons d’interpréter cette transaction ?

9 interprétations possibles

1. Les deux inputs appartiennent à Alice, qui paie 3 BTC à Bob, et récupère 1 BTC en monnaie.
2. Idem, mais 1 BTC pour Bob, 3 BTC de monnaie retourne à Alice
3. Alice fournit 1 BTC, Bob 3 BTC, et il récupère des UTXO de même montant.
4. Alice fournit 3 BTC, Bob 1 BTC, Bob récupère 3 BTC, Alice 1 BTC (= Alice a payé Bob 2 BTC).
5. Alice paie 4 BTC à Bob.
6. Alice possède tous les inputs et tous les outputs.
7. Alice possède tous les inputs, un des outputs va à Bob et l'autre à Carol.
8. Alice paie 3 BTC, Bob 1, un des outputs est à Carol, l'autre à David.
9. Alice et Bob paient 4 BTC à Carol.

Attaques hors de la block chain

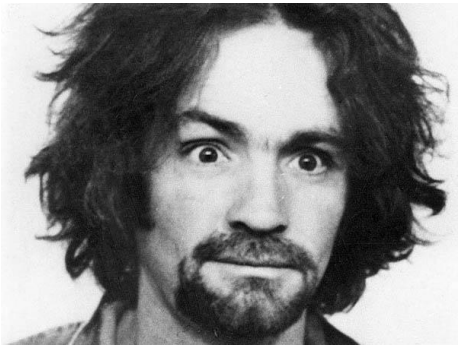
- Bitcoins achetés sur un exchange avec KYC : l'exchange sait non seulement tout de votre activité Bitcoin, mais aussi de votre identité réelle et travaille main dans la main avec le gouvernement = Game Over
- Analyse de trafic (FAI, gouvernement, hackers...) :
  - Détection de données typiques d'un logiciel Bitcoin = vous utilisez Bitcoin
  - Transactions qui partent de chez vous = vos transactions
- Custodial Wallet (pas de KYC) : la société qui fournit le service sait absolument tout de vos transactions (mais pas un utilisateur tiers)
- Client léger (Electrum) : le serveur que vous interrogez pour avoir vos transactions sait tout de vous.
- Adresse échangée de façon non sécurisée (SMS, forum…) : l'attaquant peut la lier à votre identité
- Achat en ligne : vous êtes obligé d'associer une adresse postale à votre transaction.

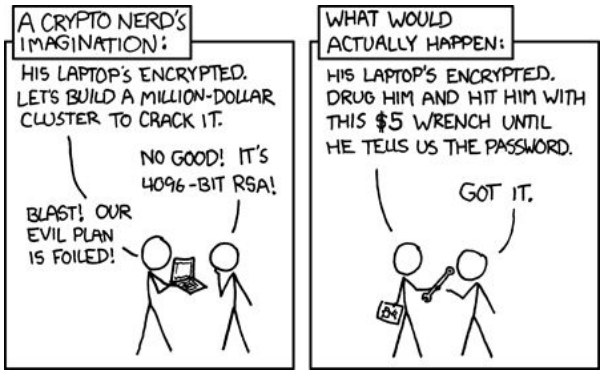
Adversaires #1: le voisin oisif

- Simples curieux, qu’ils vous connaissent personnellement ou non.
- Peu ou pas d’intentions malveillantes, faible motivation, faibles moyens.
- Ils peuvent toutefois faire beaucoup de mal en répandant des informations exploitées par des attaquants plus sérieux

Adversaires #2: la brute

- Attaquants “traditionnels”: cambrioleurs, délinquants, stalkers et psychopathes divers.
- Malveillant, but généralement assez clair, opportuniste, peu de connaissances techniques mais ne recule pas devant la violence.
- Ne recherchera pas de lui-même des informations sur vous, mais vous ciblera s’il est capable de vous identifier.
- Conséquences potentiellement dramatiques pour vous-mêmes et votre famille.





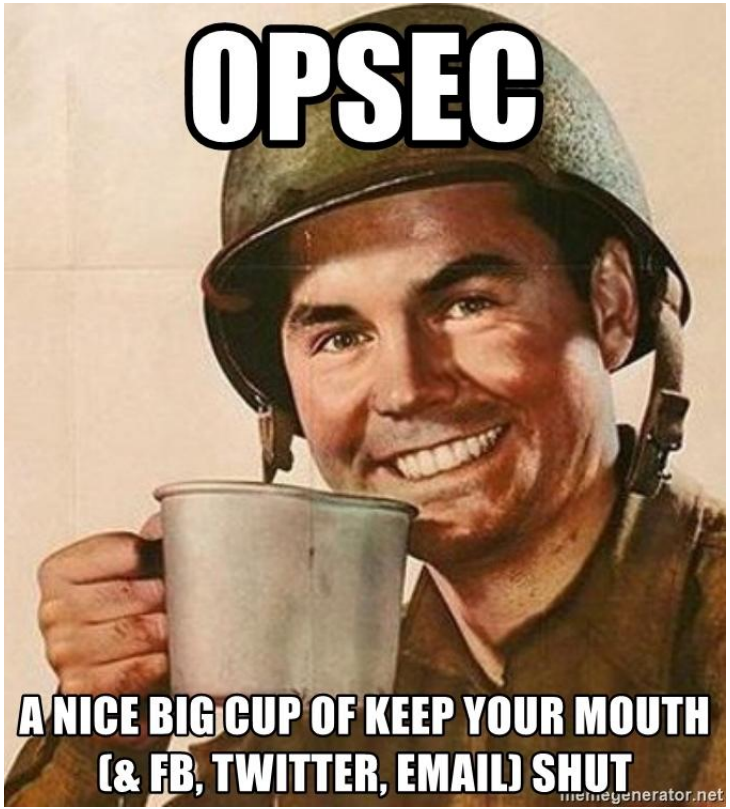
Adversaires #3: le hacker

- Cyber-délinquant indépendant (ne travaille pas pour un gouvernement).
- Proche de la brute dans ses motivations, mais mode d’attaque différent.
- Peut être totalement opportuniste (ransomware) ou vous cibler directement si vous avez une mauvaise opsec.
- Conséquences moins effrayantes qu’une agression physique, mais il peut être plus difficile de s’en défendre.

Adversaires #4: le Leviathan

- Comprend les forces de l’ordre et les agences de renseignement (motivations, compétences et moyens assez différents).
- Adversaire ultime : motivé, compétent, beaucoup de moyens financiers et légaux, en plus du monopole de la violence.

3. Comment s’en protéger ?



Quels sont les différents axes ?

1. Votre modèle de menaces : Qui ? Pourquoi ?
2. Préventif vs curatif
3. *Plausible deniability* : plus important que le secret lui-même, il faut pouvoir cacher qu'on a un secret !

Évitez de fournir votre identité pour acheter

C’est le moment de risque maximum, et aussi celui qui est le plus simple à résoudre : n’achetez JAMAIS sur un site avec un KYC.

Il existe plusieurs alternatives (plus ou moins réalistes) :

1. Achat en cash
2. Plateforme d’échanges décentralisée : BISQ
3. Faites-vous payer en bitcoins (au noir bien sûr)
4. Miner
5. Voler

Ne réutilisez JAMAIS une adresse

- Cela rend extrêmement facile le travail d'analyse (certitude absolue que le destinataire de toutes les transactions est la même personne, pareil pour les paiements !)
- Cela ne vous met pas seulement vous en danger, mais aussi tous ceux qui échangent avec vous :
  - Ceux qui paient cette adresse se retrouvent associés à votre identité
  - Ceux que vous payez ensuite avec cette adresse sont aussi contaminés

Autres conseils de bons sens

- Ne pas acheter des biens ou des services qui nécessitent de fournir des informations perso (surtout pour acheter des choses illégales)
- Ne pas rechercher une transaction qui nous concerne sur un explorateur de blocs (en tout cas pas sans Tor, cf ci-dessous)
- Anonymiser son trafic internet avec Tor, ne pas laisser traîner sa vie sur les réseaux sociaux
- Se taire, ne pas parler de ses bitcoins, de combien on en a, comment on les garde etc...

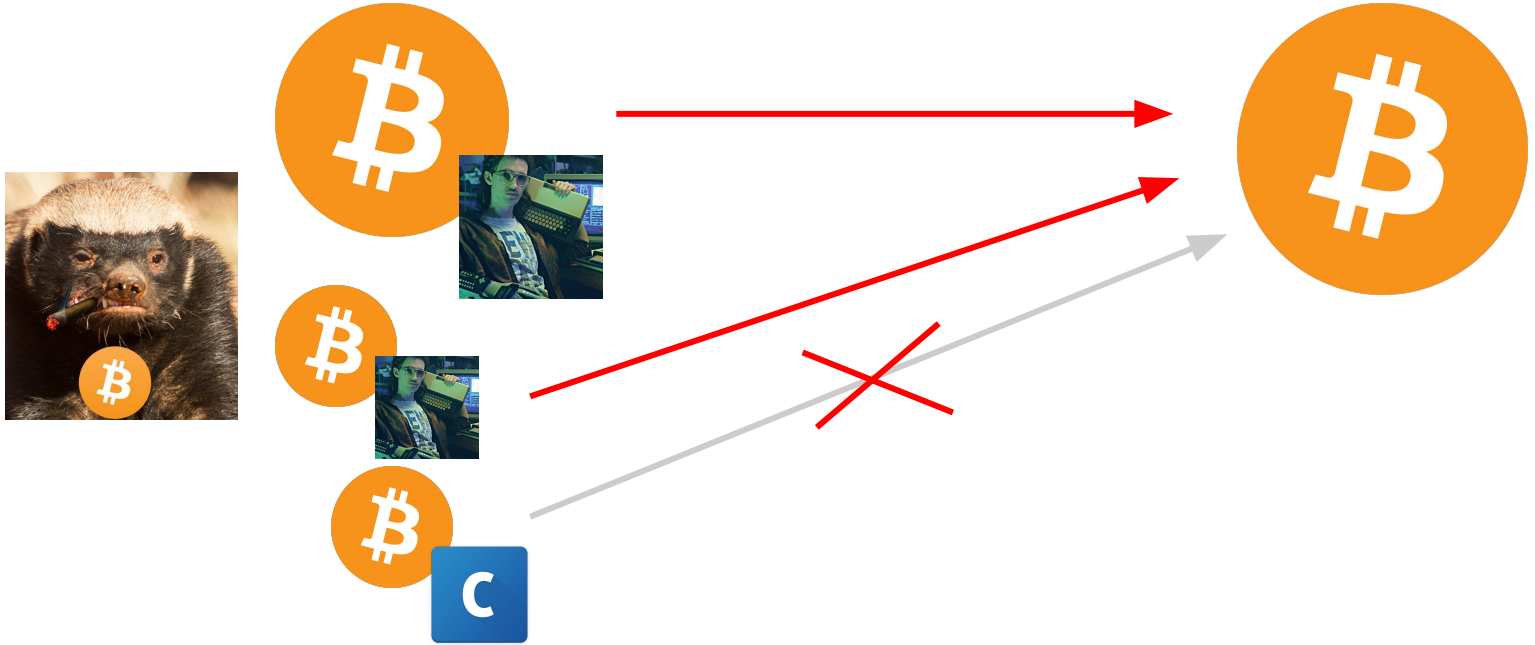
Et bien sûr, utilisez **votre propre full node**, n'utilisez jamais d'autres wallets ni de services tiers

Quand le mal est fait...

Il existe aujourd’hui un certain nombre de techniques qui permettent de contrer l’analyse de block chain. Le principe est de construire des transactions qui vont à l’encontre de l’heuristique utilisée afin de la rendre inopérante.

Coin Control

- L’idée est de séparer les UTXO qui proviennent de différentes sources pour éviter de désanonymiser bêtement des UTXO “sain”.



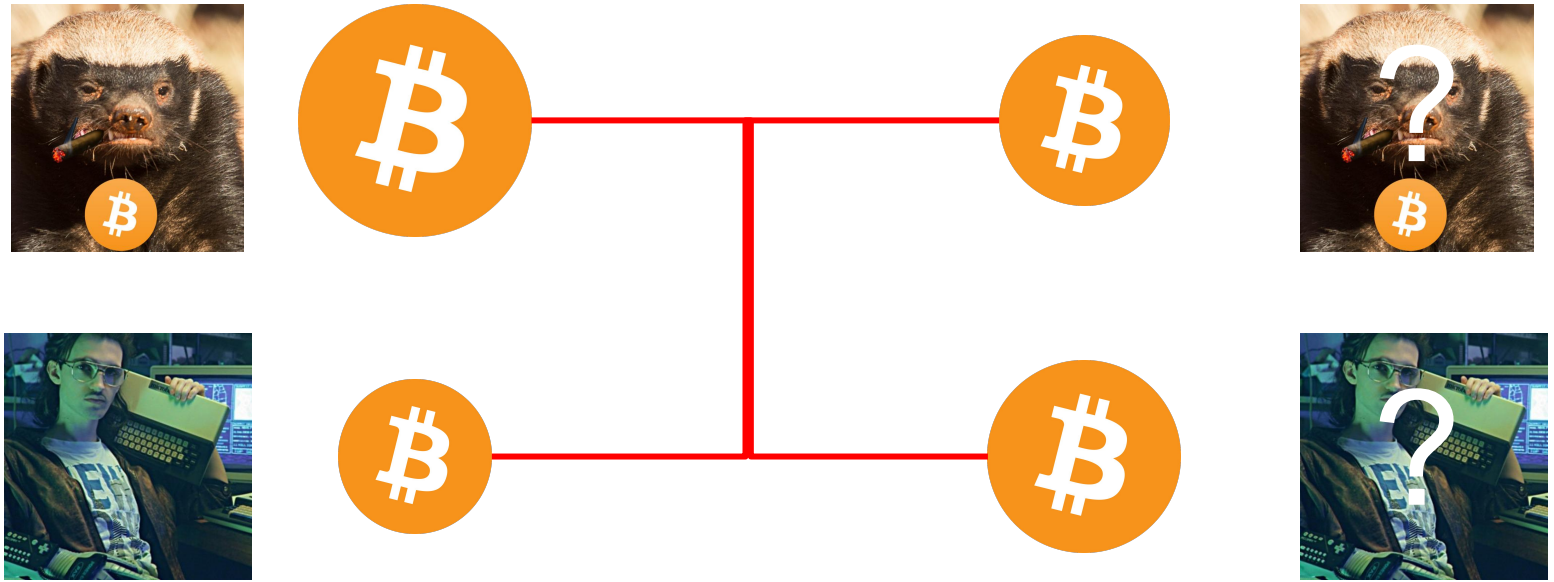
Mixer

- C’est une technique ancienne et un peu dépassée. L’utilisateur qui souhaite anonymiser ses bitcoins les envoie à un service, le “mixer”, qui, va ensuite les renvoyer sur d’autres adresses (en gardant une commission).
- Il y a de nombreux inconvénients :
  - Le mixer sait exactement quels bitcoins appartient à qui
  - Le mixer peut facilement voler les bitcoins
  - L’anonymisation ne peut marcher qu’à condition d’avoir suffisamment d’utilisateurs différents (il y a eu des cas de voleurs qui essayaient de blanchir leurs bitcoins dans des mixers, mais ils représentaient la quasi-totalité des bitcoins mixés à un instant t…)
- Enfin, il est possible d’exécuter la même technique sans passer par ces intermédiaires, en effectuant des dépôts puis des retraits sur des sites sans KYC (casinos en ligne, site d’e-commerce etc…)

CoinJoin

- Un CoinJoin vise à rendre inopérante l’hypothèse “tous les inputs appartiennent à la même personne”. Les participants apportent chacun un input et un ou plusieurs outputs et construisent une seule transaction dont ils sont aussi en possession des outputs.
- Les outputs peuvent être tous du même montant ou non. Plus il y a de participants, plus le gain de confidentialité est important. Il est recommandé d’effectuer plusieurs tours pour plus d’anonymat.
- Une transaction CoinJoin est “trustless”
- Attention à bien séparer les UTXO “sain” de ceux qui peuvent être encore marqués (cf coin control)

PayJoin (ou P2EP)



Lightning Network

- Le réseau Lightning est mieux connu comme solution de scalabilité, mais il est intéressant de remarquer qu'il est aussi **excellent d'un point de vue confidentialité**.
- À part l'émetteur et le destinataire d'un paiement, **personne ne peut savoir qui paie quoi sur Lightning !**
- Toutefois les transactions de création et de fermeture de canaux de paiement peuvent toujours être désanonymisées, car ce sont des transactions bitcoins normales.
- Des hubs de paiement pourraient voir qui ouvre des canaux vers eux, mais pas les paiements (sauf collusion entre différents hubs pour désanonymiser les utilisateurs, assez compliqué en pratique).

Sidechain

- Une autre piste est “d’ancrer” des bitcoins dans une autre block chain avec une transaction particulière.
- Tout ce qui se passe sur cette sidechain est invisible sur la block chain de Bitcoin.
- La sidechain peut opérer selon des règles différentes de celles de Bitcoin, et utiliser d’autres techniques cryptographiques qui favorisent la confidentialité, comme Confidential Transaction sur Liquid.

bisq

- Le seul exchange décentralisé :
  - pas de serveur
  - pas de dépôt (ni bitcoins, ni fiat)
  - même pas de personne morale
- P2P + Tor pour les communications
- Séquestre pour éviter les arnaques
- La plupart des paiements en fiat se font toutefois par virement bancaire

- Implémentation de CoinJoin dans un wallet léger.
- Fonction poussée de Coin Control : obligation d’étiqueter tous ses outputs.
- Le but est de passer les outputs de son wallet par plusieurs “rounds” de CoinJoin, ce qui permet d’augmenter son “anonymity set” (parmi combien d’outputs le mien est-il mélangé ?)
- Il existe aussi d’autres implémentations de CoinJoin, comme Joinmarket

Autres innovations à venir

- Plusieurs développements en cours au niveau des scripts de signature :
  - Schnorr : les multisigs deviennent indiscernables des signatures standards, mais nécessite un softfork
  - MAST : permet de cacher les branches non exécutées d'un smart contract
  - Taproot : mélange des 2 améliorations précédentes pour programmer des smart contracts plus évolués
- Confidential Transaction : déjà en prod sur la sidechain Liquid

Commandes utiles

- getblock [txid] (verbosity=2)
- getrawtransaction [txid]
- decoderawtransaction [hex]