12 比特币交易的匿名性

比特币和匿名性 Bitcoin and Anonymity。

一般我们会把匿名性和隐私性privacy联系起来。但其实比特币不是完全没有用户名字，只不过用的是化名而已。所以是pseudonymity。

其实比特币的匿名性还不如现金，因为现金上是没有任何身份信息的。所以很多非法的交易都是用现金。但是现金的问题在于不方便携带和运输。跟银行存款相比呢？比特币的匿名性要比银行账户的匿名性要好一些，因为银行账户是实名制。比特币是不需要实名制的。其实国内的银行以前不是实名制的，可以使用化名。谁有存折就可以把钱取走。如果银行用化名的话，匿名性和比特币相比哪个更好？其实是要比比特币更好的。因为比特币的账本是公开的，谁都可以查。但是银行的账本只有银行的工作人员和司法机构能查，但是普通老百姓是没有办法查看的。

比特币当中什么情况下有可能破坏匿名性？比如说我每次收款的时候都创建一个新的公私钥对，这样每次收钱都是用不同的账户来收。但是有哪些情况下可以找出不同账户之间的关联呢？比特币交易允许有多个输入和多个输出。比如说一笔交易的输入可能有两个账户add1和add2，输出也可以有两个账户add3和add4。这个时候add1和add2很有可能就是一个人的，因为同时控制了两个账户的私钥。

为了迷惑敌人，也可以故意多几个输出地址，就是把一部分钱转回给自己的账户，一部分钱转给商家。

比特币中的账户信息也可能和你生活中的真实身份产生关联。任何比特币系统和实体世界发生联系的时候，都有可能泄露真实身份。比如说资金的转入和转出。一开始的时候你没有比特币，就需要用现实的钱去买，这个时候有资金流入，就很有可能查到谁买的这个比特币。很多国家都有反洗钱法。追溯洗钱最好的办法就是追溯比特币的转入和转出。如果有特别大的资金，转入和转出如果不想引起司法机构的注意其实是很难的。

用比特币做支付的时候，也有可能泄露自己的真实身份。比如说买咖啡的时候用比特币支付。但是这是个bad idea。因为交易延迟很长，而且交易费太高。而且这个时候就会建立支付账户和你的真实身份的联系。

举一个例子，曾经一个信用卡机构，将信用卡号码取哈希之后将所有的交易记录都公开了用于科研。结果有的科研发现，我们可以通过现实生活中观察到的交易经过很少次数的筛查之后就可以精准定位信用卡的哈希值和真实身份之间的联系。说明信用卡记录是不应该公开的。

但是比特币的交易记录都是公开的。比如说去买咖啡，不仅咖啡店能知道你的账户和身份关联，所有知道或者观察到你在这个时间用比特币购买咖啡的人都会知道你的账户和身份关联。

中本聪的匿名性保持的最好。因为从来不花比特币。一旦花了比特币就要和真实世界产生联系。他宁愿这些比特币一个都不花，也不愿意暴露身份。所以他不贪财哈哈哈哈。

Silk Road。 eBay for illegal drugs。买的都是违禁品。为了逃避司法制裁，交易用比特币。底层网络用洋葱路由。并且在美国使用匿名寄送。老板在最后在旧金山被捕了。被抓的时候被没收了十几万比特币。但是一个都没有花，因为一花就暴露身份了。被抓的具体细节没有披露，但是据说是因为他用同一台电脑登陆了真实身份的社交网站，所以被发现了。他被抓了之后，有人又开了第二版的silk road。比特币的匿名性并没有我们想象的好。中本聪能长期保持匿名是个例，而且他不干什么坏事。比特币的项目走上正轨之后，中本聪就消失了。功成身退。中本聪干这个事情不是为了钱，也不是为了名。

Hide your identity from whom?

普通老板姓还是很容易的。但是像丝路那种的，FBI想抓你还是很容易的。

假设你是一个比特币用户，能采取什么样的方法尽量的提高自己的匿名性?

application layer

---

network layer(P2P)

Network layer如何实现匿名性？比如说你去网吧，用网名发布一个帖子，别人有办法知道是你发布的帖子么？ 你去网吧得有身份证，所以只需要查一下你的ip地址，和你的机位对应起来，就知道你是谁。所以比特币要实现匿名性，必须先实现网络层的匿名性。否则就可以从ip推出你在物理世界中的身份。但网络层的匿名性比较好解决，可以用多重转发。比如说洋葱路由。每一个包并不是由sender直接发给receiver的。而是中间经过好多个节点。每个节点都只知道自己的上一个节点是谁，而不知道这个包最早是由哪个节点发出来的。中间的这些节点中有可能是坏的，但只要有一个节点是好的，就可以隐藏最初发件人的身份。

Application layer怎么实现匿名性？所有的币都可以追溯到最初的来源。最简单的办法就是把不同的人的币混在一起，就是Coin Mixing。不光是区块链，在所有需要达成匿名性的领域都可以这样。就是把你的身份信息和别人的混在一起，从而分不清楚谁是谁的。这个地方我们可以把币都混在一起，从而分不清楚这些币都是哪里来的。但是如何实现呢？有一些专门做coin mixing的网站。你可以把币存进去，他们随机打乱之后，退回给你。往往退回给你的币就不会你最初存进去的币。但是通过不同的analysis可以分析出来。加上现在的币圈没有信誉很高的做coin mixing的机构。因为往往coin mixing都是匿名的，卷币跑路的话就没有办法了。很多在线钱包都自带coin mixing的属性。往往你存进去的钱，花的时候不一定是来源于都一个地方。但是在线钱包不一定保证要履行coin mixing的功能。还有一般的比特币交易所，会有天然的coin mixing性质。比如说你上传一些比特币上传到交易所，过一段时间卖了。过一段时间又买入以太币，再换成莱特币，最后换成比特币。这个时候的比特币就不是最一开始的比特币了。但这个前提是交易所不会泄露相关的交易信息。

为什么保护隐私性难度很大？ 因为区块链本身是公开的，而且是不可篡改的。实际上不可篡改性对匿名性是灾难性的，因为一旦泄露，就没有办法挽回了。而且用户之间的关联性很致命，如果一个泄露出去了，就全部都泄露出去了。

零知识证明Zero Knowledge Proof：

零知识证明是指一方（证明者）向另外一方（验证者）证明一个陈述时正确的，而无需透露除该陈述是正确的外的任何信息。

比如说证明一个比特币账户是我的，怎么证明?

显然我不能泄露私钥，我可以签名。然后你可以通过我的公钥验证我的签名。这个过程中没有泄露私钥。但是这个例子是不是零知识证明是存在争议的，因为我虽然没有泄露私钥，但是泄露了这个私钥产生的一个签名。这个和零知识证明的定义是有一定差距的。

同态隐藏：E是加密函数

1. 如果x，y不同，那么他们的加密函数值E(x)和E(y)也不相同
2. 给定E(x)的值，很难反推出x的值。
3. 给定E(x)和E(y)的值，我们可以很容易地计算出某些关于x,y的加密函数值。

   - 同态加法：通过`E(x)`和`E(y)`计算出`E(x+y)`的值
   - 同态乘法：通过`E(x)`和`E(y)`计算出`E(xy)`的值
   - 扩展到多项式

例子：

Alice向Bob证明她知道一组数x和y使得x+y=7，同时不让Bob知道x和y的具体数值。

简单版本：

Alice把E(x)和E(y)的数值发给Bob。Bob通过收到的E(x)和E(y)计算出来E(x+y)的值（利用性质三），Bob同时计算E(7)的值，如果E(x+y)=E(7)（利用性质1），那么验证通过，否则验证失败。

但在这个例子中，Bob可以使用brute force，遍历所有的x的input，猜到x是多少。

更复杂的版本：

可以对x和y分别做一些随机处理，但是保证x和y的加和是不变的。

有什么办法可以在中心化的系统中，既然央行充当查验交易合法性的角色，又不让央行知道每个账户的具体信息？那就是虚拟货币的编号不能是央行产生的，而是我自己产生的。而这个编号不能告诉央行。但是不告诉央行的问题就是央行没有办法旅行查验交易合法性的角色了。

这个时候我们就有盲签方法：

用户A提供SerialNum，银行在不知道SerialNum内容的情况下返回签名Token，减少A的存款。

用户A把SerialNum（明文）和Token交给B完成交易。

用户B拿SerialNum（明文）和Token给银行验证（因为银行需要检验double spending），银行验证通过，增加B的存款。

银行无法把A和B联系起来。

中心化

这样设计的目的就是A自己产生一个编号（银行不知道），然后银行给A一个币。然后A拿着这个币去给B花，B把A告诉自己的编号和银行的签名给银行进行验证，因为银行不知道当初A让他签名时候的编号，所以银行此时也没办法知道B手里的币的来源是哪里。

零币和零钞：

比特币一定程度上是匿名性的，但是不能消除账户之间的关联性。那么我们有没有办法来设计一种加密货币从一开始就消除用户之间的关联性?

零币zerocoin和零钞zerocash在协议层就融合了匿名化处理，其匿名属性来自密码学保证。

零币系统中存在基础币（可以是比特币）和零币，通过基础币和零币来回转换，消除旧地址和新地址的关联性，其原理类似于混币服务。（花的时候，用基础币换成零币，花零币的时候只需要用零知识证明证明你花的币是一个系统中合法的币就可以了，而不知道具体是哪一个币）

零钞系统使用zk-SNARKs协议，不依赖一种基础币，区块链中只记录交易的存在性和矿工用来验证系统正常运行所需要关键属性的证明。区块链上既不显示交易地址，也不显示交易金额，所有交易通过零知识验证的方式进行。

但零币和零钞都不是主流的加密货币。一是他们为了匿名性付出了一定的代价，因为数学原理复杂，导致性能缩减。二是可能需要强匿名性的用户没有那么多，毕竟不干坏事。三是虽然零币和零钞虽然在数学原理上提供了很强的匿名性，但也不是百分之百的匿名性，因为还会和实体发生交互。