比特币的匿名性(bitcoin and anonymity)

什么叫匿名？一般来说，匿名是跟隐私保护联系在一起的。比特币中不要求用真名，可以用公钥产生的地址，所以比特币具有一定的匿名性。也就是你可以产生任意多的地址，然后用不同的地址干不同的事情。它用的是化名，但它不是完全没有名字，所以有人把它称为preudonymity。

比特币和银行存款哪个匿名性更好? 银行账户是实名制，你得提交身份信息，然后才能注册银行账户，而比特币不需要，从这点上看比特币匿名性要好。

其实，以前银行是可以用化名的，如果银行账户匿名的话隐私性和匿名性与比特币相比哪个好? 银行账户更好。比特币区块链的账本是公开的，所有人都能查到，每个人都可以上区块链把整个信息下载下来。而银行的账本是受控制的，银行的工作人员可以查到，一些司法手段也可以调取银行的信息，但普通百姓是查不到别人的账的。

比特币系统中什么情况下有可能破坏匿名性? 有的人推荐每次收款都用一个新的地址，这样的话可以有不同的地址，谁也不知道哪些是属于你的，看起来好像匿名性很强，但实际上这些地址是可以被关联在一起的。

比如网上购物，比特币交易允许有多个输入多个输出。而多个输入有可能是同一个人，因为这个人可能同时控制了这两个账户的私钥。为什么要有两个输入:因为你买的东西很可能不是你某个账户上全部的币。有多个输出，很可能有一个是找零钱的地址。这种交易一般都是比特币钱包软件生成的。很多交易软件每次交易的时候都会生成一个新的找零的地址，也是为了隐私保护。

有没有可能把输入地址和输出地址也关联起来呢？比特币生成交易的时候并没有规定找零钱的地址在outputs中出现的位置，所以想知道哪个是找零的地址也并不容易，但有些情况下可以分析出来。

比如第一个地址账户上有4个比特币，第二个有5个。产生的两个输出第一个输出转入6个比特币，第二个输出转入3个比特币。那很明显转入3个比特币的输出是找零的，因为如果它是商家的地址，就用不着两个inputs，任意一个输入都比3大。通过这种方法我们可以把输入地址和输出地址也关联起来。

所以，一个比特币用户可以生成很多个账户，但这些不同的账户是可以被关联起来的。 匿名缺陷2是比特币交易中总会有与现实关联的地方，比如场外交易，比如在交易所里注册后进行资金的转入转出。反洗钱的一个常用手段就是盯住资金的转入转出链。 。

所以回到前面的问题:比特币的匿名性有多好?匿名是跟隐私保护相关联的，但问题在于:你不想向谁暴露身份(hide your identity from whom)？如果你不想让身边的亲戚朋友知道，这是比较容易实现的。如果是非法组织，从事黑市活动，那保护起来就难多了。

一个比特币用户能采用什么样的方法尽量提高个人的匿名性?以前曾讲过，**比特币系统是运行于应用层(application layer)的，底层是(network layer)。所以要提高匿名性可以从两个方面入手。**

①网络层怎么提高匿名性? 在现实中，如果一个人去网吧发了帖子，别人是有办法知道他是谁的。因为他的身份证代表了他的身份，这和他的IP地址是有很大关联性的。 而网络层的匿名性是比较好解决的。区块链是个新生事物，但网络层的匿名性学术界已经有了很好的方案:**多路径转发。跟洋葱路由(TOR)是一样的原**理。即消息不是由发出者直接发送给接收者，中间要经过很多次转发。中间的每一个节点，只知道它的上一个节点是谁，但并不知道最早发出消息的人是谁。当然中间一些节点可能是坏的，但路径上只要有一个节点是诚实的，就能够把最初发起人的身份隐藏起来。这也是**洋葱路由**的基本原理。

②应用层怎么提高匿名性? 把不同人的币混在一起(coin mixing)，即把你的身份跟别人的身份混在一起，让别人分不清楚谁是谁。不光是区块链，在其他各个需要匿名的领域都能用到。有一些专门做coin mixing的网站，

提供一定的服务收取一定的服务费。所有想做coin mixing的人把币发给网站，网站内部进行一些重组，然后你再把币取回来，这时取出的币就不是发布到网站上的币了，它是随机抽取一些币给你。

coin mixing真正实施起来有一定的复杂性，如果设计不好的话，别人可以根据你当初存进去币的数额，推断出来哪些币是你存进去的。 而且，在当今的区块链的世界里，没有什么信誉度非常高的coin mixing的服务。很多coin mixing的服务它本身也是要保持匿名的，它匿名的后果是:有可能投进去的币被他卷款跑路了，投币者是一点办法都没有的。

实际上并不一定非要做coin mixing，有一些应用本身也带有coin mixing的性质，比如在线钱包。很多人会把钱存入**在线钱包**里，在线钱包就会把这些人的币混起来，再取回自己的币时可能就不是当初存进去的币了。但在线钱包并不保证要履行coin mixing的功能。

还可以通过**加密货币的交易所**，交易所一般有天然的coin mixing的性质。前提是交易所不会泄露提币、存币的记录，否则也是不行的。

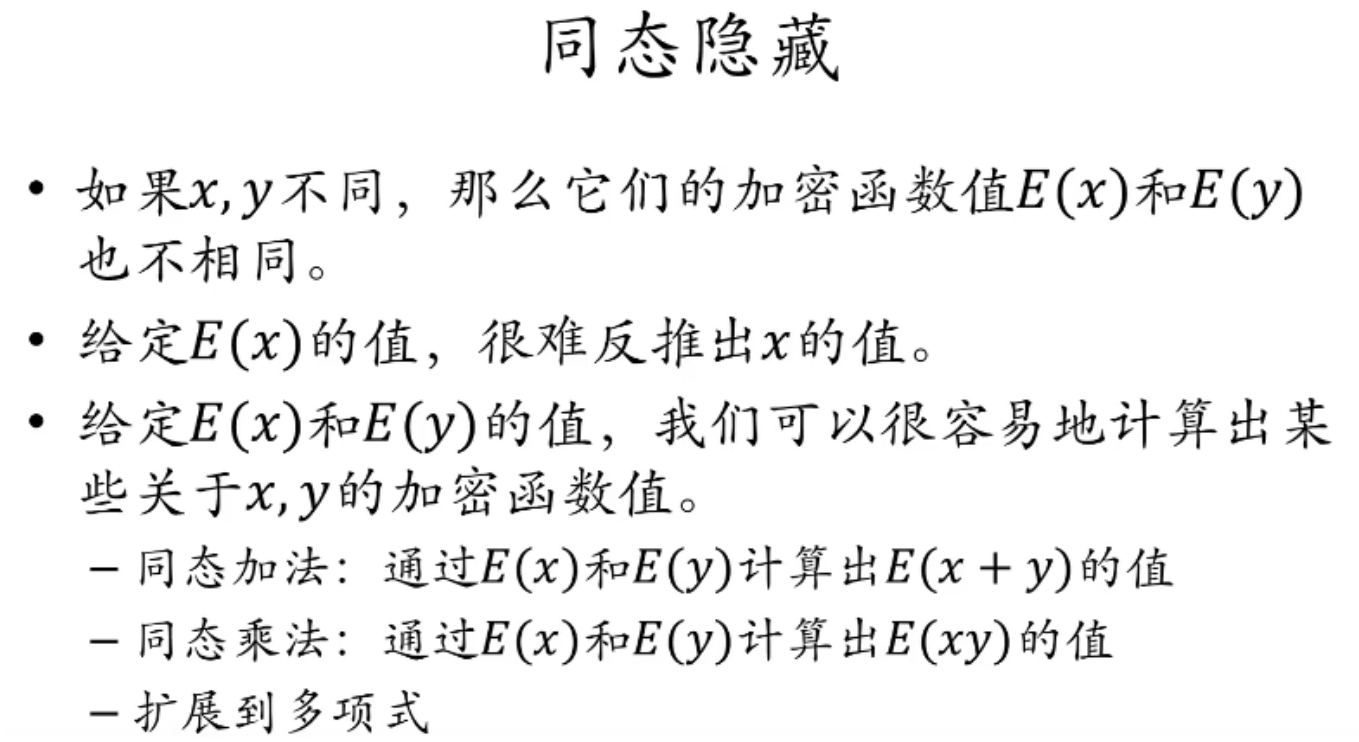
为什么保护隐私性难度挺大？本质原因是区块链是公开的，而且是不可篡改的。不可篡改性对于隐私保护来说是灾难性的。

零知识证明:

概念如图(第44分第50秒)所示:零知识证明是指一方(证明者)向另一方(验证者)证明一个陈述是正确的，而无需透露除该陈述是正确的外的任何信息。

例如:要证明一个账户是我的，只需要我给出私钥就行。但私钥不能直接泄露，所以就给出由私钥产生的签名，假设对方是知道这个账户的公钥的，那么就可以验证签名的正确性。这是不是一个零知识证明其实是有争议的，因为我给出了私钥之外的其他信息，具体算不算要看应用场合。

同态隐藏:



零知识证明的数学基础是同态隐藏。 如图(第48分第25秒)是同态隐藏的性质。 第一个性质说明加密函数值E不会出现碰撞collision free，这跟哈希函数有所不同，哈希函数是可能出现碰撞的collision resistance。这个性质反过来说明如果E(x)和E(y)是相等的，那么x、y也是相等的。(该语句是上面语句的逆否命题)

第二个性质说明加密函数是不可逆的，知道加密后的值，没办法推出加密前的值。hiding

第三个性质是最重要的，叫作同态运算。它说的是对加密之后的函数值进行某些代数运算，等价于对这些输入直接进行代数运算然后再加密。 同态加法:加密值的和等于和的加密。 同态乘法:加密值的乘积等于积的加密。

举一个例子:如图(第51分第53秒)所示Alice想要向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。

简单的解答版本如图(第53分第27秒) ·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E(x+y)的值(利用了性质3) ·Bob同时计算E(7)的值，如果E(x+y ）=E(7)，那么验证通过，否则验证失败。

Bob可以用蛮力算法，一个一个试而计算出x和y的值，因此Alice要对x和y的值做一些随机化处理，保证x和y加起来还是不变的。

不考虑去中心化的前提下，前面在讲double spending时讲过，要对每一个数字货币进行编号就能防止double spending。回到这节课讲的隐私保护问题，央行是什么都知道的，那么有没有什么办法让央行做中心化的记账检测double spending，又不让它知道呢?即虚拟货币的编号不能是央行产生的，改成自己产生的，又不会被篡改掉。

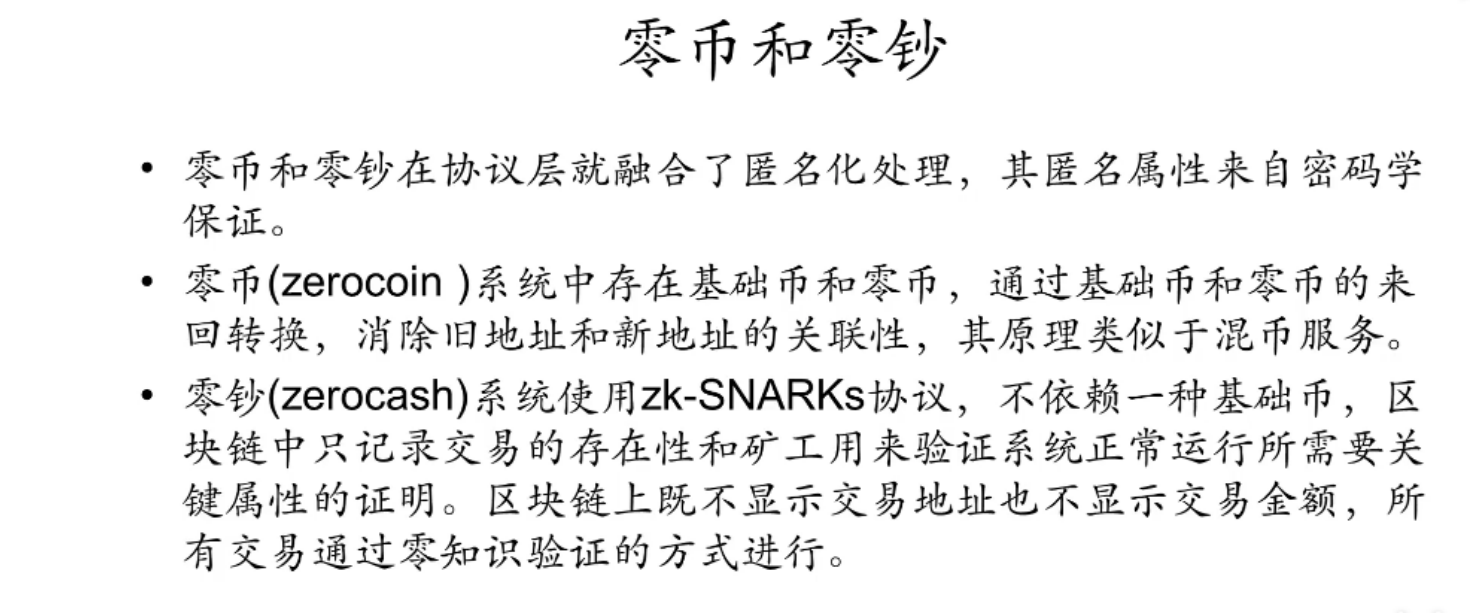
这里就要用到盲签方法。如图(第59分第39秒)。 ·用户A提供SerialNum，银行在不知道SerialNum的情况下返回签名Token，减少A的存款 ·用户A把SerialNum和Token交给B完成交易 ·用户B拿SerialNum和Token给银行验证银行验证通过，增加B的存款 ·银行无法把A和B联系起来 ·中心化

解读如下: ·用户A提供序号，银行进行签名但此时看不到序号的内容，A要取钱所以银行要减少A的存款。 ·A给B转账交易的时候把序号和签名给B，这个时候序号是明文，B是可以看到序号的具体内容的。 ·B把序号和签名给银行验证，这个时候序号也是明文，这一步验证的目的是检测double spending。 这样设计的好处是:银行不知道B的币是从哪来的。

零币和零钞(它们也是加密货币，跟比特币是一类属性):

比特币在很大程度上提供了匿名性，但它不能完全消除关联性，那么我们能不能设计一种新的加密货币，这个货币从一开始的结构设计上就用了密码学的原理保证了匿名性，所以就有了零币和零钞。

如图(第62分第47秒)零币和零钞 ·零币和零钞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(zerocoin)系统中存在基础币和零币，通过基础币和零币的来回转换，消除旧地址和新地址的关联性，其原理类似于混币服务。 ·零钞(zerocash)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。



这是专门为匿名性设计的加密货币。 零币中存在基础币(比如比特币)和零币。用的时候要证明本来是有一个基础币，让基础币变得不能花费(unspendable)，然后换取一个零币，零币在花的时候只需要用零知识证明你花掉的币是系统中存在的某一个合法的币就行了，但是不用透露你花的是系统中具体的哪一个币。这是跟比特币的一个本质区别，比特币是每一笔转账交易都要说明币的来源。这样才能证明花的币的真实性不是凭空捏造出来的。 但零币和零钞不是这样，零币和零钞是说证明的时候可以从数据上保证你花的币是以前区块链上某个合法存在的币，但不知道具体是哪个。这样的话就把关联性破坏掉了，就没法追溯了。

零钞没有基础币，是完全的零币。零钞和零币也不是100%匿名安全的，在影响匿名安全的因素中依然有一个因素无法解决，就是与实体发生交互的时候。比如有人想拿这些币干坏事，把很大的金额转换成这种加密货币的时候，或者是把这些加密货币转换成现金的时候，仍然要暴露身份。这些加密货币数学上设计的再好，只是说对已经在区块链当中的转账有匿名性，跟外界交互的匿名性仍然是一个弱点。所以它依然无法提供100%的匿名。