假期准备学习工作总结——第四周

本周的主要学习内容有**密码学Hash函数**、**英语**和**区块链理论** 。

在之前吕老师的线上授课时，带领大家梳理了一下区块链学习的知识框架，密码学中的Hash函数是BTC系统的基石，虽然也提到了椭圆曲线密码学，但由于数学基础薄弱且搁置的时间有点久，所以先进行比较熟悉的Hash函数的学习，然后再去学习数论的相关内容。在学习本章Hash函数时，涉及到了一些概率运算和逻辑运算，学习过程相对没那么顺利，在分析简单Hash的安全性问题时更是如此，因此学习进度缓慢。在学习网课和读相关著作时有很多的英文部分，所以也在坚持用APP背单词。

密码学Hash函数部分学习到的知识点总结如下：**1.**Hash函数H（）将可变长度的数据块M作为输入，产生固定长度的Hash值 h=H（M）**2.** 一个好的Hash应该具有这样的特点：对于大的输入集合使用该函数，产生的输出结果均匀分布且看起来随机。**3.** Hash函数的首要目标是保证数据完整性，对于输入任何一位或者几位的改变都将极大可能改变其Hash值。**4.** 消息认证是用来验证消息完整性的一种机制或服务，当Hash函数用于提供消息认证功能时，Hash值被叫做消息摘要。与消息认证应用类似，对于Hash函数的另外一重要应用就是数字签名，在进行数字签名过程中使用用户的私钥加密消息的Hash值，而公钥是公开的，其他任何知道该用户公钥的人都能通过数字签名来验证消息的完整性，有效抵御了攻击者想要篡改消息的行为。**5.** 如果一个Hash函数满足图1.2中的前五个要求则称其为弱Hash函数，如果第六个性质抗强碰撞性也满足，则称其为强Hash函数。一个函数如果是抗强碰撞的那么也同时是抗弱碰撞的，但反之则不一定成立。一个函数可以是抗强碰撞的，但不一定是抗原象攻击的，反之亦然。一个函数可以是抗弱碰撞的，但不一定是抗原象攻击的，反之亦然。**6.** 同加密算法一样，对于Hash函数的攻击也分为两类：穷举攻击和密码分析。**7.** 对于原象攻击和第二原象攻击，攻击者对给定的Hash值h，试图找到满足H（y）=h的y，对于m位的Hash值，穷举规模约为2的m次幂。对于碰撞攻击，攻击者试图找到两个消息或数据块x和y，满足H（x）=H（y），与原象攻击和第二原象攻击相比，其穷举规模相对较小，如果在均匀分布的0到N-1的范围内选择随机整数变量，那么在根号下N次选择后发生重复的概论就会超过0.5，因此，对于m位的Hash值，如果随机选择数据块，预计在根号下2的m次幂此尝试后就可以找到两个具有相同Hash值的数据块。

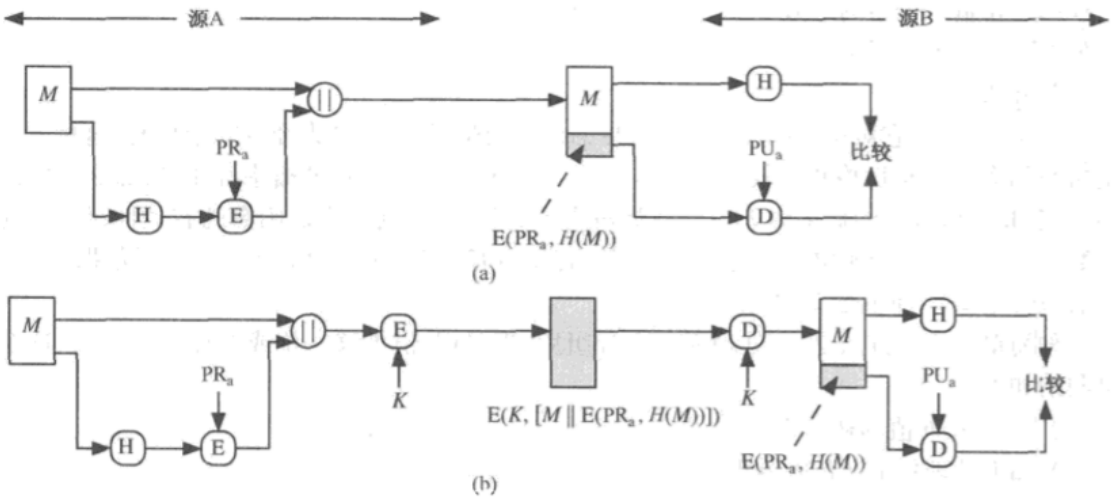

图1.1　数字签名简要示例

需　求	描　述
输入长度可变	H 可应用于任意大小的数据块
输出长度固定	H 产生定长的输出
效率	对任意给定的 x，计算 $H(x)$ 比较容易，用硬件和软件均可实现
抗原像攻击（单向性）	对任意给定的 Hash 码 h，找到满足 $H(y)=h$ 的 y 在计算上是不可行的
抗第二原像攻击（抗弱碰撞性）	对任何给定的分块 x，找到满足 $y\neq x$ 且 $H(x)=H(y)$ 的 y 在计算上是不可行的
抗碰撞攻击（抗强碰撞性）	找到任何满足 $H(x)=H(y)$ 的偶对 (x,y) 在计算上是不可行的
伪随机性	H 的输出满足伪随机性测试标准

图1.2 密码学Hash函数H的安全性需求

区块链内容的学习形式仍是接续之前的网课。总结的知识点如下：**1.** 在BTC系统中，区块链像一个去中心化的账本，比特币使用的是基于交易的这种账本模式(transaction【交易】-based ledger【账本】)。系统当中并不会显示每个账户有多少钱。**2.** 比特币系统的全节点要维护一个叫UTXO(unspent transaction output)即"还没有被花出去的交易的输出"的数据结构。区块链上有很多交易，有些交易的输出可能已经被花掉，有些还没有被花掉。而没有被花掉的输出的集合就叫做UTXO。**3.** UTXO的作用是为了检测double spending。即检测新发布的交易是否合法。因此全节点要在内存中维护UTXO这样一个数据结构，以便快速检测double spending。**4.** 每个交易可以有多个输入，也可以有多个输出，所有输入金额之和要等于输出金额之和。即total inputs=total outputs。因此一个交易可能来自多个地址，可能有多个签名。**5.** 比特币系统设计了第二个激励机制:交易费(transaction fee)。也就是你把我的交易打包在区块里，我给你一些小费。交易费一般很小，也有一些简单的交易没有交易费。**6.** 除了比特币这种基于交易的模式，与之对应的还有以太坊使用的基于账户的模式(account-based ledger)。在这种模式中，系统是要显示的记录每个账户上有多少币。比特币基于交易的模式，隐私保护性较好。缺点是比特币当中的转账交易要说明币的来源，而基于账户的模式就不用。