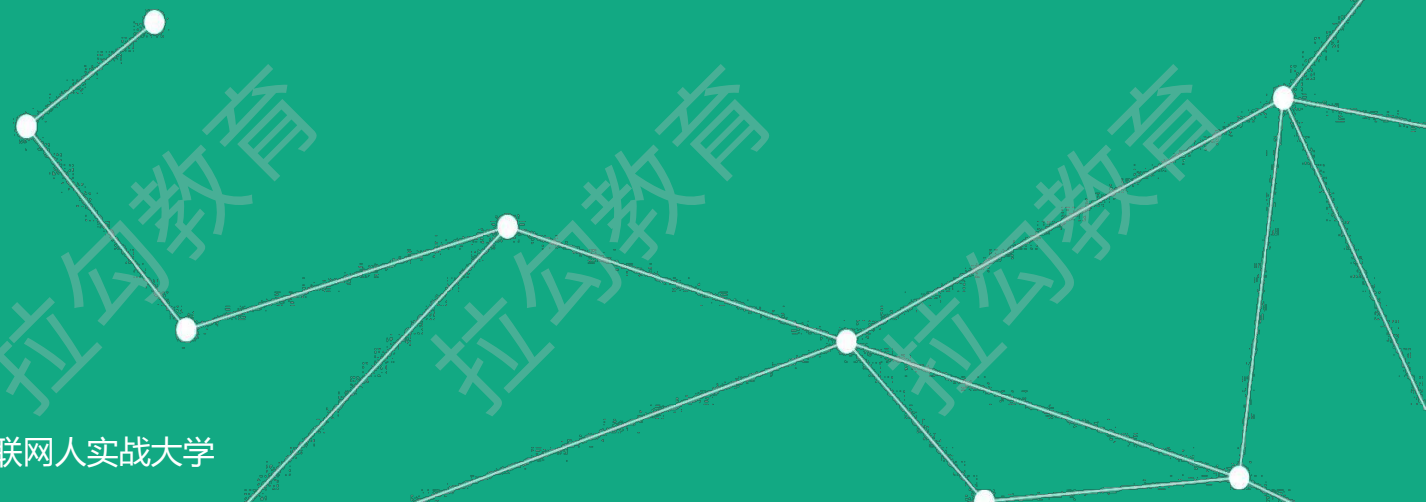
课时 06

哈希函数在 GitHub 和比特币中的应用

1. 加密哈希函数
2. SHA-1 加密算法
3. GitHub 面临的问题
4. 比特币的本质



加密哈希函数

加密哈希函数（Cryptographic Hash Function）：一个哈希函数如果能够被安全地应用在密码学中

加密哈希函数

“数字摘要” 也是通过加密哈希函数，由任意长度的一个信息转换出来的一个固定长度的哈希值

用于检验一段数据或者一个文件的完整性（Integrity）

当这个数据文件里面的任何一点内容被修改之后，通过哈希函数所产生的哈希值也就不一样了

从而就可以判定这个数据文件是被修改过的文件，也称这样的哈希值为检验和（Checksum）

加密哈希函数

常见的加密哈希函数算法：

- MD（Message Digest）算法：通过输入产生一个 128 位的哈希值出来，用于确保信息传输的完整性
- SHA（Secure Hash Algorithm）算法：常见的有 SHA-1、SHA-256 算法等，也是可以通过输入而产生一个 160 位或者 256 位的哈希值

SHA-1 加密算法

2017年

SHA-1 加密算法被正式宣布攻破了

这意味着什么呢？

那些采用 SHA-1 加密算法去验证数据完整性的应用

有可能会被人为地制造哈希碰撞而遭到攻击

SHA-1 加密算法

Git 采用 SHA-1 加密算法来做数据完整性验证

提交代码，运行“git commit”命令

Git 会将所有的这些文件，外加一些元数据（Metadata）再做一次 SHA-1 运算来得到一个新的哈希值

这些元数据里就包括了上一次 commit 时所生成的哈希值

GitHub 面临的问题

著名的代码软件托管平台 GitHub 其实也面临着同样的问题

根据 2017 年所公布的实验结果

真的要人为的去制造一个 SHA-1 哈希冲突攻击的话，现阶段的代价是非常昂贵的

比方说需要耗费 6500 年的单核 CPU 计算时间，或者说需要消耗 110 年的单核 GPU 计算时间

所以单单靠着暴力枚举的方法是不太可行的

GitHub 面临的问题

根据 Github.com 报告

一些针对 Github.com 的碰撞攻击其实是运用了一些特殊的技巧来减少这些运算时间

而这些攻击里面都会有一个固定的“字节模式”可循

所以 GitHub.com 会针对每一个上传的文件都执行一种 SHA-1 碰撞的检测

而他们所用的检测工具也是开源的（https://github.com/cr-marcstevens/sha1collisiondetection）

Linux 和 Git 之父 Linus的邮件内容：（ https://marc.info/?l=git&m=148787047422954

比特币的本质

比特币是由一个网名为“中本聪”的人所提出的在 2009 年诞生的一个虚拟加密货币

本质思想是以区块链为基础而搭建起来的一个去中心化的记账系统

所有的交易记录都存放在了一个叫区块（Block）的数据结构里面，可以看作是链表数据结构中的一个节点

当用户需要将新的交易记录打包的时候，可以自己创建一个新的区块出来，放在整个区块链的结尾

也就相当于在一个链表的结尾插入一个新的节点，而在整个区块链中的第一个区块，也就是链表的头节点

叫做创世区块（Genesis Block）

比特币的本质

区块链采用了哈希值的方式去寻找节点

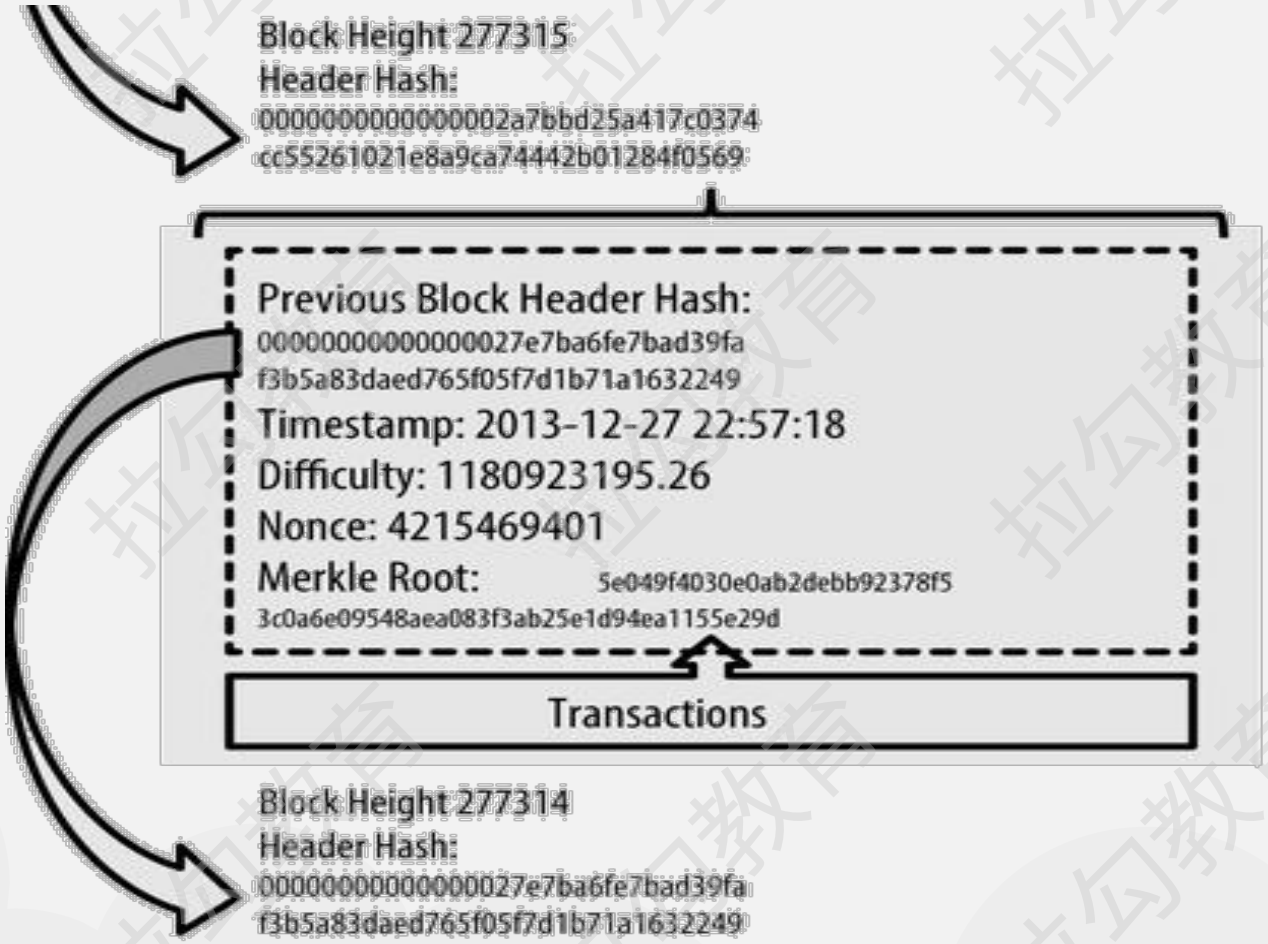
在比特币里，它采用的是 SHA-256 这种加密哈希函数，将每一个区块都计算出一个 256 位的哈希值

在每一个新的区块中都会保存着上一个区块所计算出来的哈希值，通过这个哈希值，就可以找到哪一个区块是这个新区块的上一个区块

所有的区块都可以通过这种机制去寻找上一个区块，从而遍历整个区块链，直到找到创世区块为止

比特币的本质

比特币的本质


Block Height 277315
Header Hash:
0000000000000002a7bbd25a417c0374
cc55261021e8a9ca74442b01284f0569
Previous Block Header Hash:
0000000000000027e7ba6fe7bad39fa
f3b5a83daed765f05f7d1b71a1632249
Timestamp: 2013-12-27 22:57:18
Difficulty: 1180923195.26
Nonce: 4215469401
Merkle Root: 5e049f4030e0ab2debb92378f5
3c0a6e09548aea083f3ab25e1d94ea1155e29d
Transactions
Block Height 277314
Header Hash:
0000000000000027e7ba6fe7bad39fa
f3b5a83daed765f05f7d1b71a1632249

比特币的本质