数字货币和区块链 - 密码学（1）

山东大学网络空间安全学院

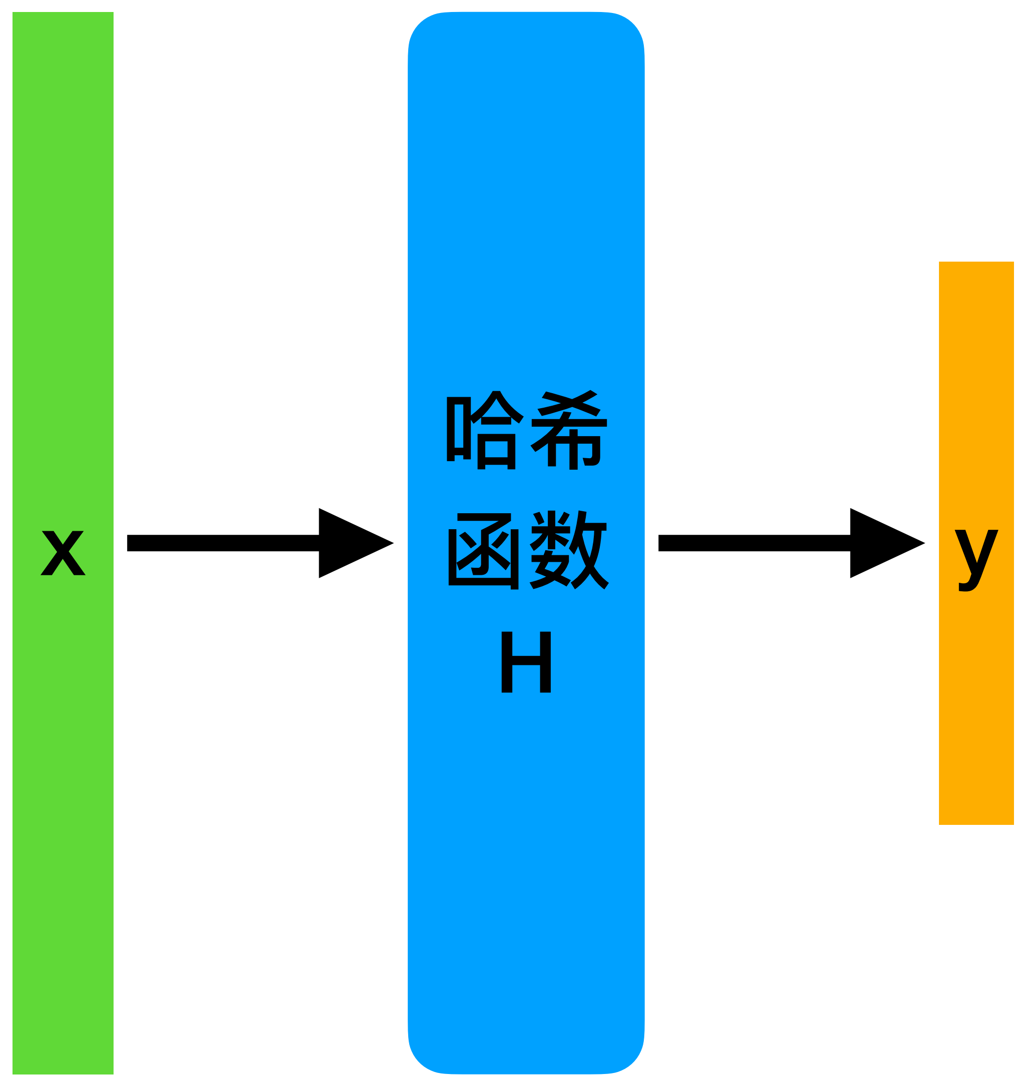
钱宸 2023年10月30日

密码学温故知新

- 哈希函数
- 零知识证明
- 签名
- 盲签名

哈希函数

- 哈希函数（杂凑函数、散列函数）是计算机科学中的一个重要概念
- 将很长的数据压缩成短的哈希值 y=H(x)



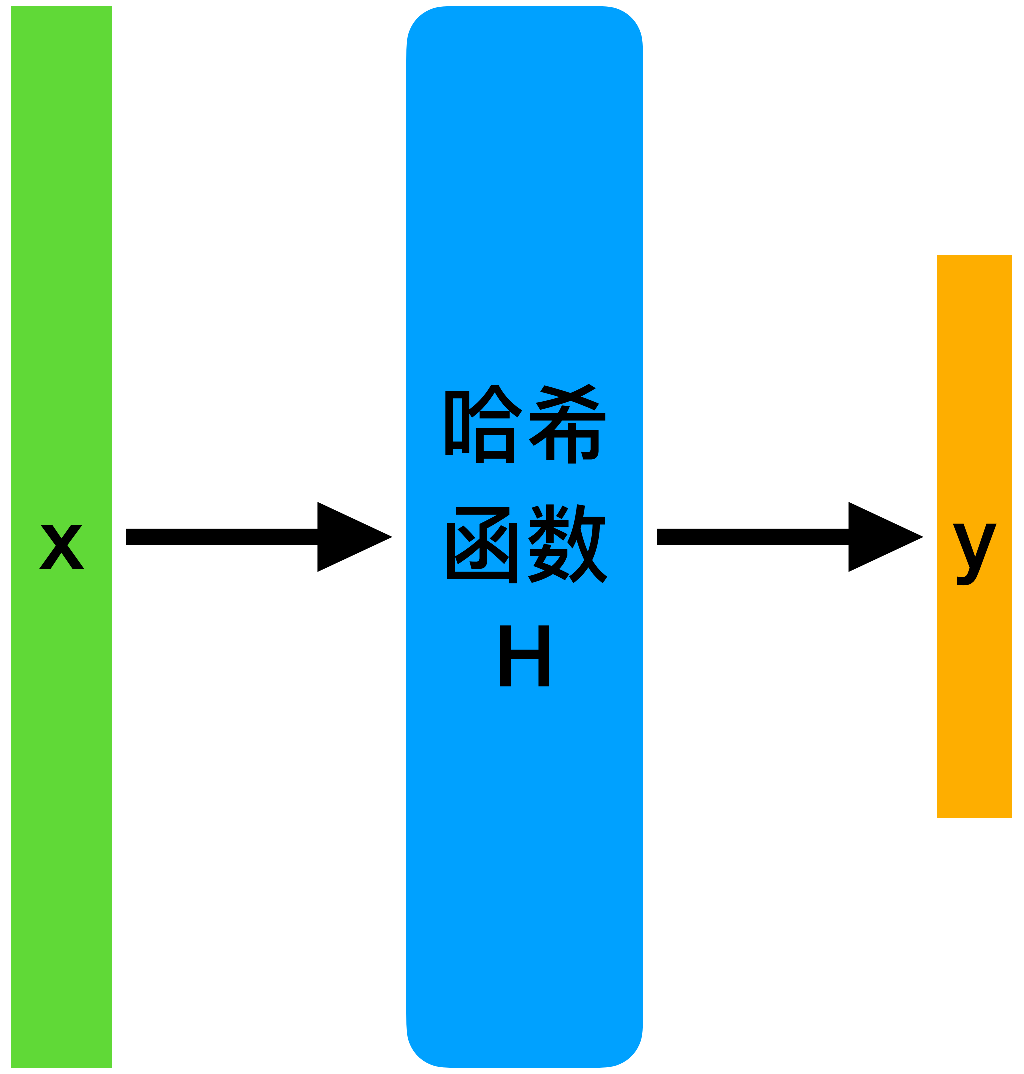
哈希函数的作用

- 哈希函数被广泛应用于多种算法当中：
  - 哈希表 →高效查找
  - 资料保护
  - 语音识别特征
  - 数据传输完整性

什么是哈希函数?

- ~~将很长的数据压缩成短的哈希值 y=H(x)，其中|y| < |x|~~
- 反例：$\mathrm{H}(x) = x_1$，该函数截取输入的第一位。
  - |y|<|x| 但是输出并不能代表输入

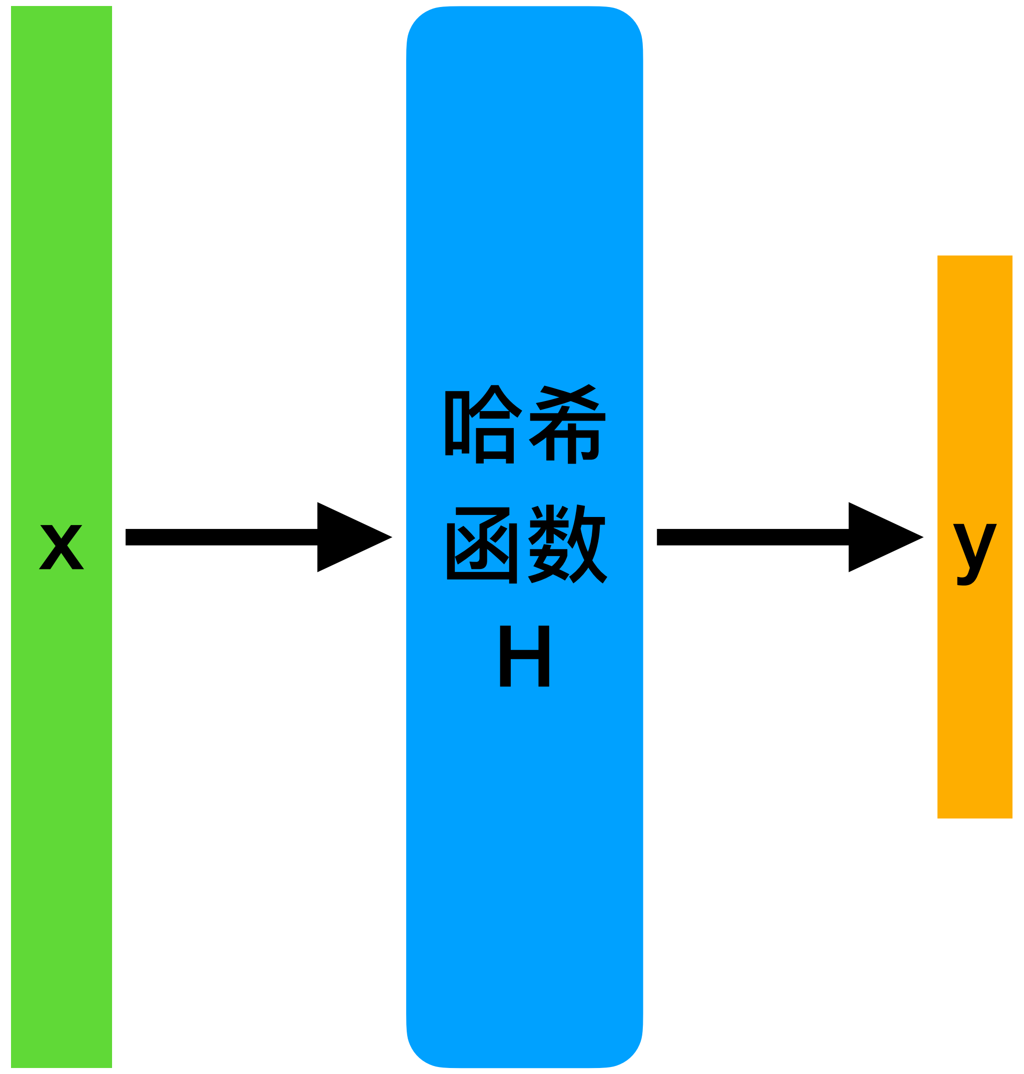
什么是哈希函数?



- 哈希函数所希望达到的效果：
  - 输出能用较短的比特作为输入的摘要
  - 函数H为哈希函数当且仅当
    - $\mathsf{H}: \{0,1\}^m \to \{0,1\}^n, m > n$
    - $\forall y \ \nexists x_1, x_2 . \mathsf{H}(x_1) = \mathsf{H}(x_2) = y \wedge x_1 \neq x_2$

$\{0,1\}^m \rightarrow \{0,1\}^n, m > n$ 鸽巢原理，H为多对一映射。

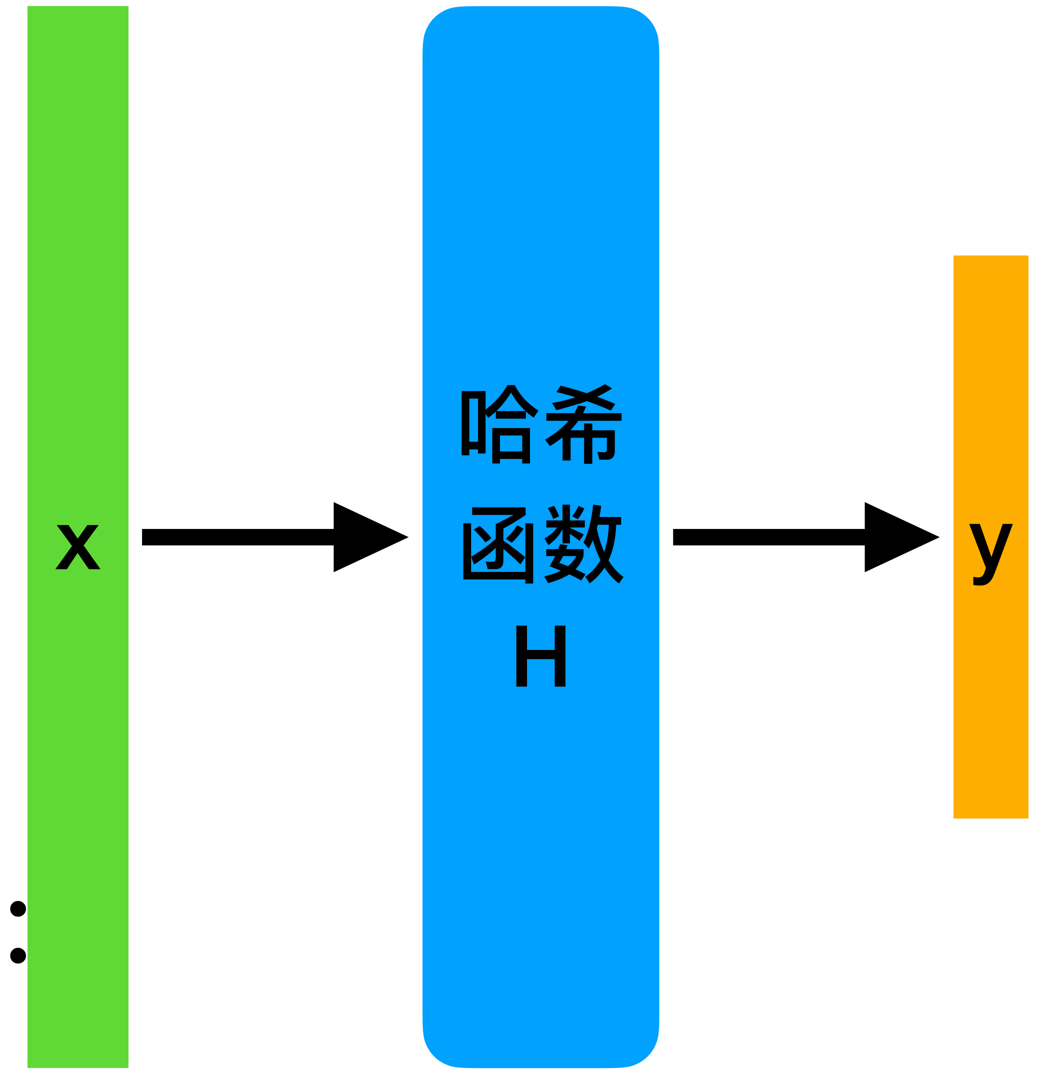
什么是哈希函数?



- 虽然H是多对一的映射，但是多对一的输入很难找?
- 函数H为哈希函数当且仅当
  - $\mathrm{H}:\{0,1\}^m \to \{0,1\}^n, m > n$
  - $\forall \mathscr{A} \in \mathrm{PPT}\,.\,\Pr[\mathrm{H}(x_1) = \mathrm{H}(x_2) \wedge x_1 \neq x_2 \mid \mathscr{A} \to (x_1, x_2)] = \mathrm{negl}$
  - 其中 PPT表示所有多项式大小概率图灵机，negl表示可忽略函数。

鸽巢原理，存在这样的x_1, x_2满足条件。构造常数函数$\mathscr{A}:() \to (x_1, x_2)$。

什么是哈希函数?



- （抗碰撞）哈希函数定义：
  - 哈希函数是一组函数确定的函数族$\mathscr{H} = \{H_k\}_{k \in \{0,1\}^l}$：
    - $\forall k \in \{0,1\}^l . \mathrm{H}_k : \{0,1\}^m \rightarrow \{0,1\}^n, m > n$
    - $\forall \mathscr{A} \in \mathrm{PPT}$.
      $\Pr[\mathrm{H}_k(x_1) = \mathrm{H}_k(x_2) \wedge x_1 \neq x_2 \;\; k \xleftarrow{\$} \{0,1\}^l; \mathscr{A}(H_k) \rightarrow (x_1, x_2)] = \mathrm{negl}$
    - 其中 PPT表示所有多项式大小概率图灵机，negl表示可忽略函数。

为了简化文字，通常用含密钥的哈希函数H_k来表示

历史中的哈希函数

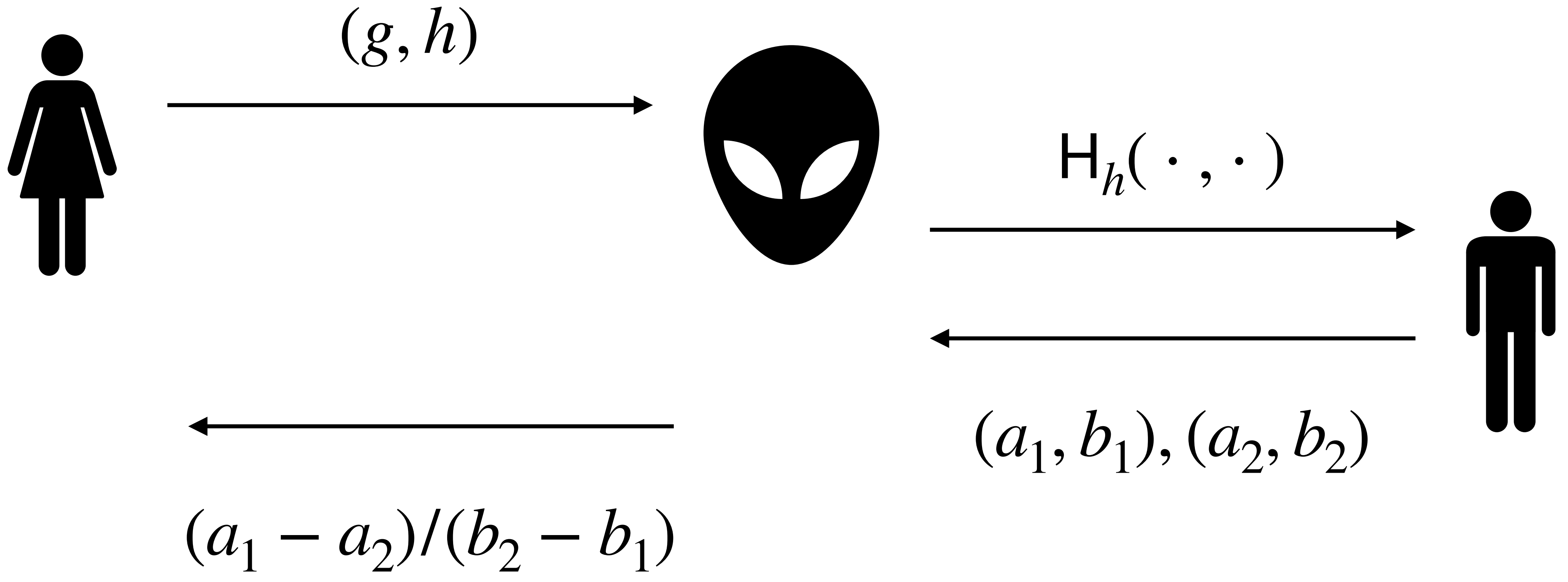
- 历史上MD5和SHA-1是使用最多的哈希函数
- 2004年MD5的首个碰撞攻击被王小云院士发现：
  - M1:0THdAsXm7sRpPZoGmK/5XC/
KtYcSRn6rQARYPrj7f4lVrTQGCfSzAoPkiIMlcUFaCFEl6PfN
  - M2:0THdAsXm7sRpPZoGmK/5XC/
KtQcSRn6rQARYPrj7f4lVrTQGCfSzAoPkiIMl8UFaCFEl6PfN
- M1和M2的BASE64解码以后会得到相同的MD5值。
- SHA-1哈希函数在2005年也被王小云院士等人发现能在2^{63}步内找到碰撞
- 2012年10月，NIST选择了KECCAK作为SHA-3新标准

另外一些哈希函数的构造方式

- MD5和SHA-1都是基于对称密码方式构造，是否真正存在哈希函数?
- 假设有个素数阶群$\mathbb{G}$，其中离散对数是难的：
  - $(g, g^x) \to x$ 是困难的
  - 令$\mathrm{H}_h(a, b) := g^a \cdot h^b$
    - $(a, b) \in \mathbb{Z}_p^2, g^a \cdot h^b \in \mathbb{G}$
    - 为什么$\mathrm{H}_h(a, b)$很难找到碰撞?

另外一些哈希函数的构造方式

- $\mathsf{H}_h(a, b) := g^a \cdot h^b$
  - 令$h = g^x, a_1 + b_1 \cdot x = a_2 + b_2 \cdot x$



- 如果存在$\mathscr{B}$能够找到该哈希函数的碰撞，则存在$\mathscr{C}$能够解决$\mathbb{G}$上的离散对数问题。

密码学温故知新

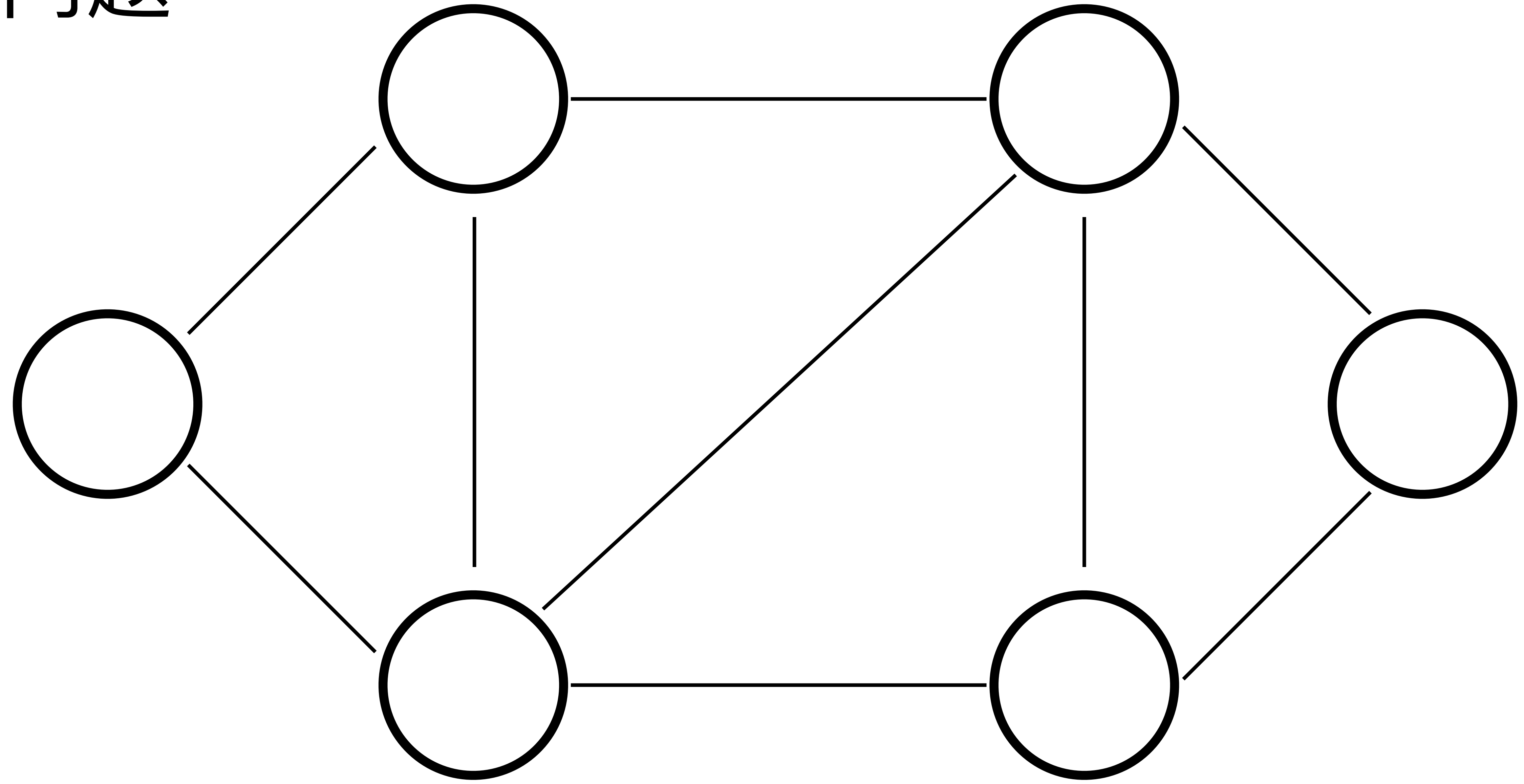
- 哈希函数
- 零知识证明
- 签名
- 盲签名

零知识证明

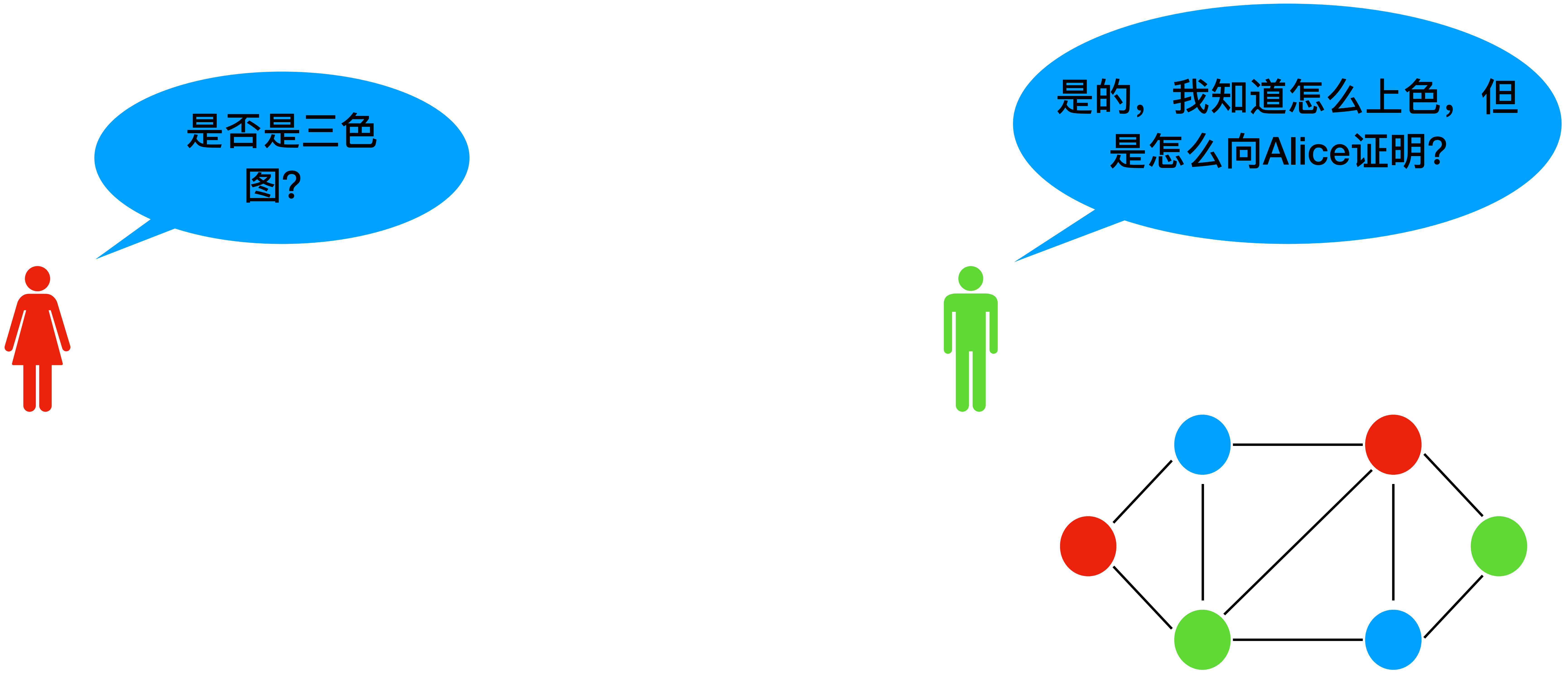
- 数字货币基于信用→如何在证明价值的时候保证匿名性?
- 零知识证明：
  - 在证明一个事件的同时，不泄露任何其他的信息。
  - 简单的例子：三色问题。

零知识证明 - 三色问题

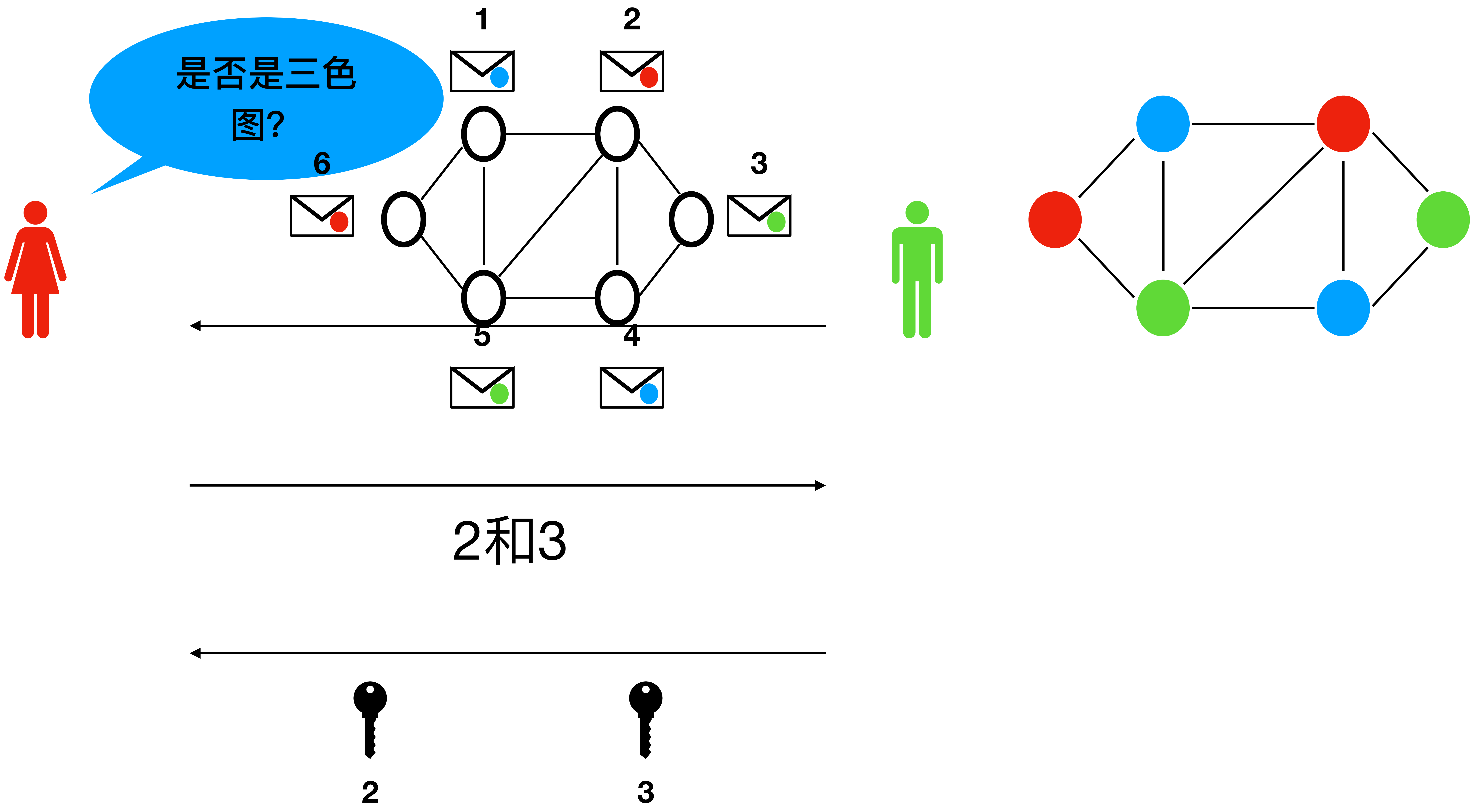
- 三色问题：
  - 给定一个无向图$G = (V, E)$，是否存在一个用三种不同的颜色给G的顶点上色的方式使得任意两个相邻的顶点颜色都不同。
  - NP-Complete问题



零知识证明 - 三色问题



零知识证明 - 三色问题



零知识证明 - 三色问题

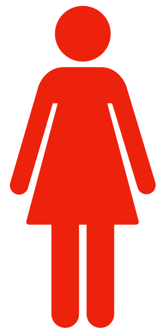
- 几个要点：
  - 每次Bob证明的时候都要重新排列所有颜色
  - 信封必须保证，里面的内容不能被篡改（密码学中用Commitment承诺实现）
  - Alice每次都只能验证一组边上的两个顶点颜色不一样

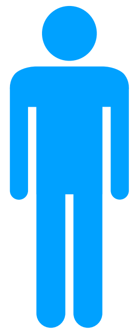
零知识证明 - 离散对数

- 如何利用类似的想法（随机应答的方式）证明数学难题?
- 假设有个素数阶群$\mathbb{G}$，其中离散对数是难的：
  - $(g, g^x) \to x$ 是困难的
- 如何证明我知道$f = g^x$?

零知识证明 - 离散对数

- 如何证明我知道$f = g^x$?





$$r \xleftarrow{\$} \mathbb{Z}_p \qquad \xrightarrow{\quad R = g^r \quad}$$

$$\xleftarrow{\quad h \quad} \qquad h \xleftarrow{\$} \mathbb{Z}_p$$

$$z = xh + r \qquad \xrightarrow{\quad z \quad} \qquad g^z = f^h \cdot R$$

零知识证明 - 离散对数

- 为什么能证明$f = g^x$?
- 假设有一个$R = g^r$同时对于两个不同的h_1, h_2都可以应答z_1, z_2
- 那么 $x = (z_1 - z_2)/(h_1 - h_2)$

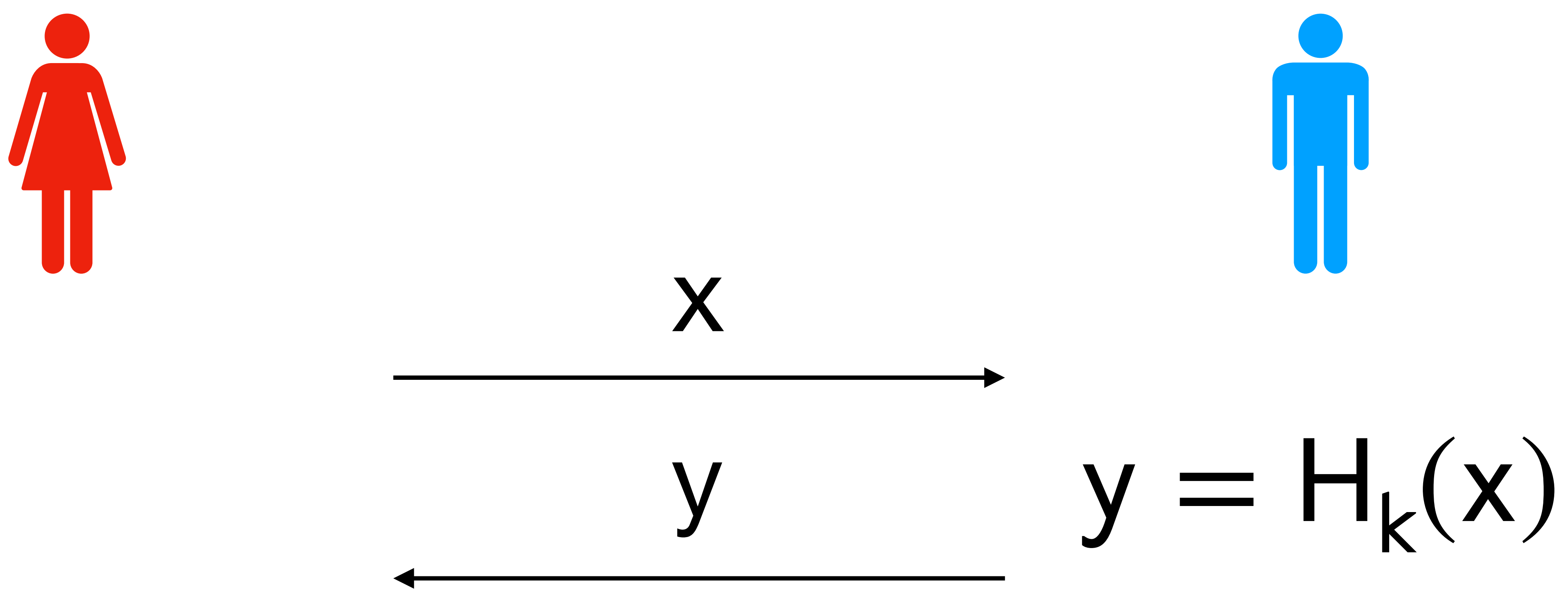
零知识证明 - 从交互到非交互

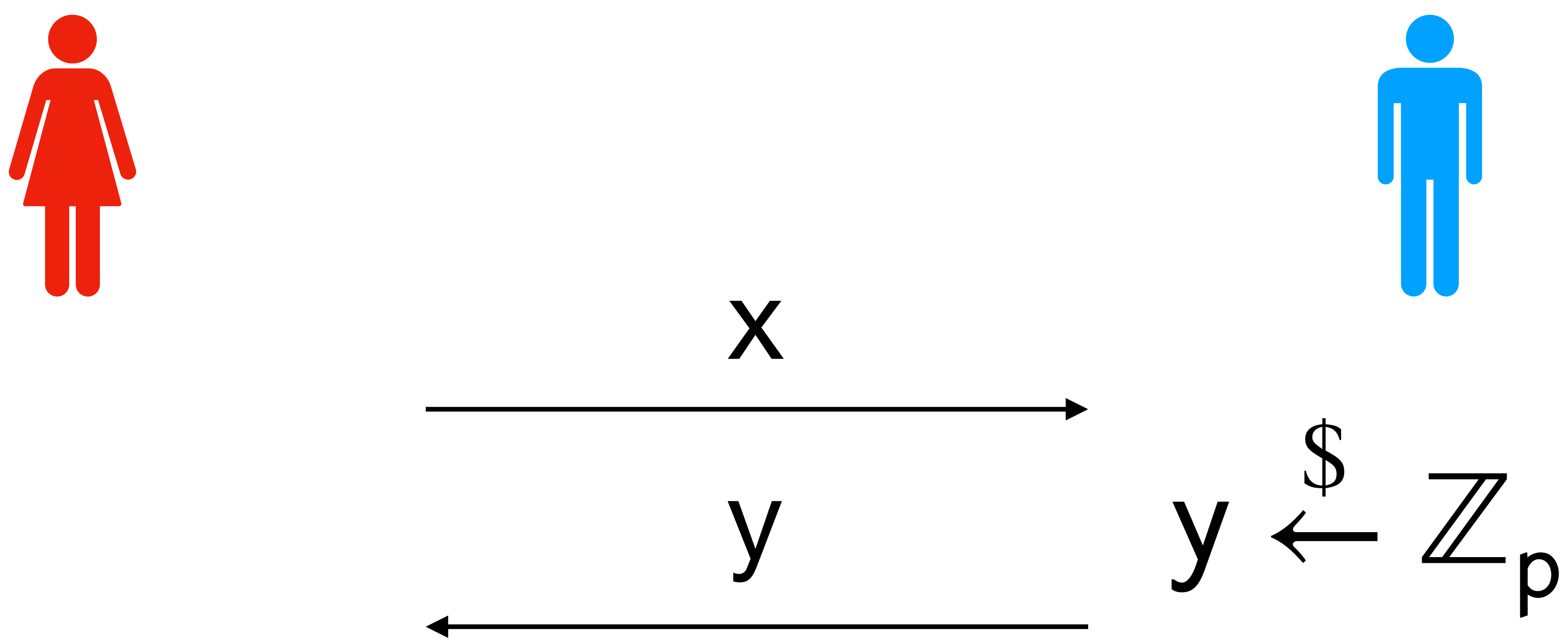
- 观察离散对数的证明过程
- Bob产生的h仅是一个随机生成的元素
- 完美的哈希函数也能产生随机元素$h = \mathrm{H}(f, g, R)$
  - 可以去除交互的过程
  - $R = g^r; h = \mathrm{H}(f, g, R); z = xh + r$
  - Alice生成$\pi = (R, h, z)$
- 这个过程可以被更广泛地应用Fiat-Shamir转化

补充：完美的哈希函数 - 随机寓言机

- 之前我们有提到（抗碰撞的）哈希函数
- 但是，哈希函数还有不同的性质
  - 弱一些的：单向性（给定输出找不到输入），思考：为什么更弱呢?
  - 强一些的：完全随机性$I(X;Y) = H(X) - H(X \quad Y) = H(Y) - H(Y \quad X) = 0$
- 零知识证明中需要完全随机性
  - 随机寓言机模型（Random Oracle Model）
  - 理想模型 - 实际并不存在！但是大幅方便了证明。

随机寓言机模型





零知识证明 - 非交互式证明

- 一个非交互式的零知识证明由(Setup, Prove, Verif)三个算法组成
- Setup$(1^\lambda) \rightarrow$ crs
- Prove(CRS, x, w) $\rightarrow \pi$
- Verif(CRS, x, π) $\rightarrow \{0, 1\}$
- 正确性：正常产生的证明都能被验证
- 知识证明：能从证明中提取出所证明的信息（不严谨表述)
- 零知识：证明中不包含额外的信息（不严谨表述)

非交互式证明

- 要证明我知道x，使得$f = g^x$
- 输出证明：$\pi = (R = g^r, h = \mathsf{H}(f, g, R), z = xh + r)$
- 验证：$g^z = f^h \cdot R$
- 在随机寓言机模型中该证明是零知识的！
- 第一步：将哈希函数转化成为随机寓言机$\mathcal{O}$
- 第二步：提前选取一个随机的h，在证明的过程中将随机寓言机替换为$\mathcal{O}'$
  - $\mathcal{O}$和$\mathcal{O}'$在所有输入上都相同除了，$\mathcal{O}'(f, g, R) = h$