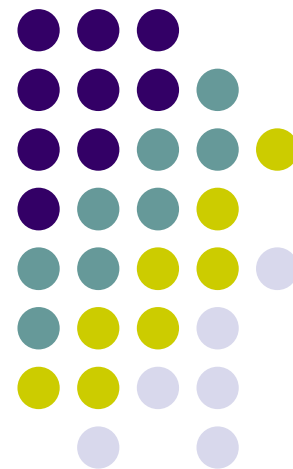


密码学与复杂性理论

林东岱

中国科学院软件研究所
信息安全国家重点实验室

2010年2月3日





信息安全问题

机密性

- 怎样保护不想让别人知道的一些信息，小的可以是私人日记、信用卡信息、银行密码，大的可以是政治、军事或外交机密，如导弹的布署等。

完整性

- 消息的接收者能够验证消息在传送过程中没有被修改或者替换，完整性保证了发送方发送的消息和接收方接收到的消息的一致性。

可认证性与零知识性

- 包括身份认证和数据源认证，用于确定人、计算机和文件的真实身份，以防假冒、替换或否认。

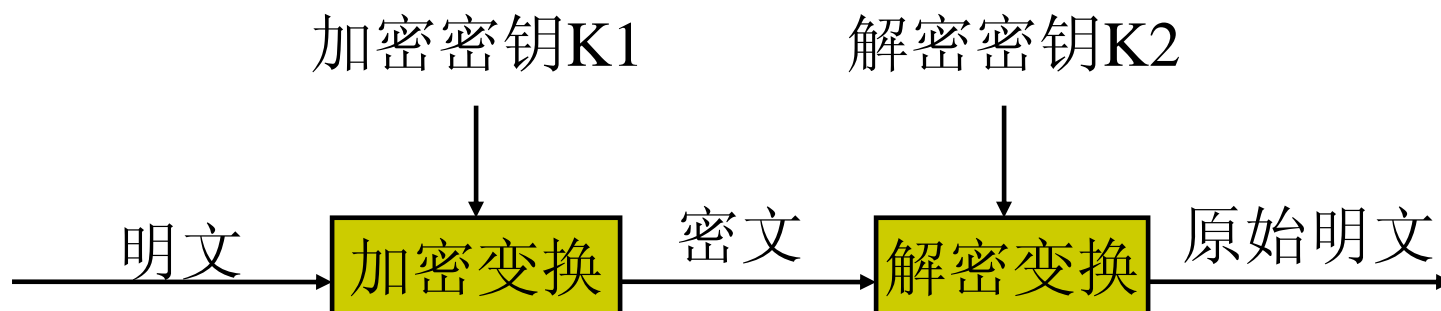


什么是密码学？

- Cryptography = hiding + writing
- 密码学是研究保密通信的一门科学。它研究在不安全的环境中，如何把所要传输的信息在发给接收者之前进行秘密转换以防止第三者对信息的窃取。
- 密码学是研究设计与分析克服敌手的不良影响的协议的一门科学

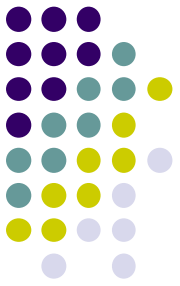


算法和密钥

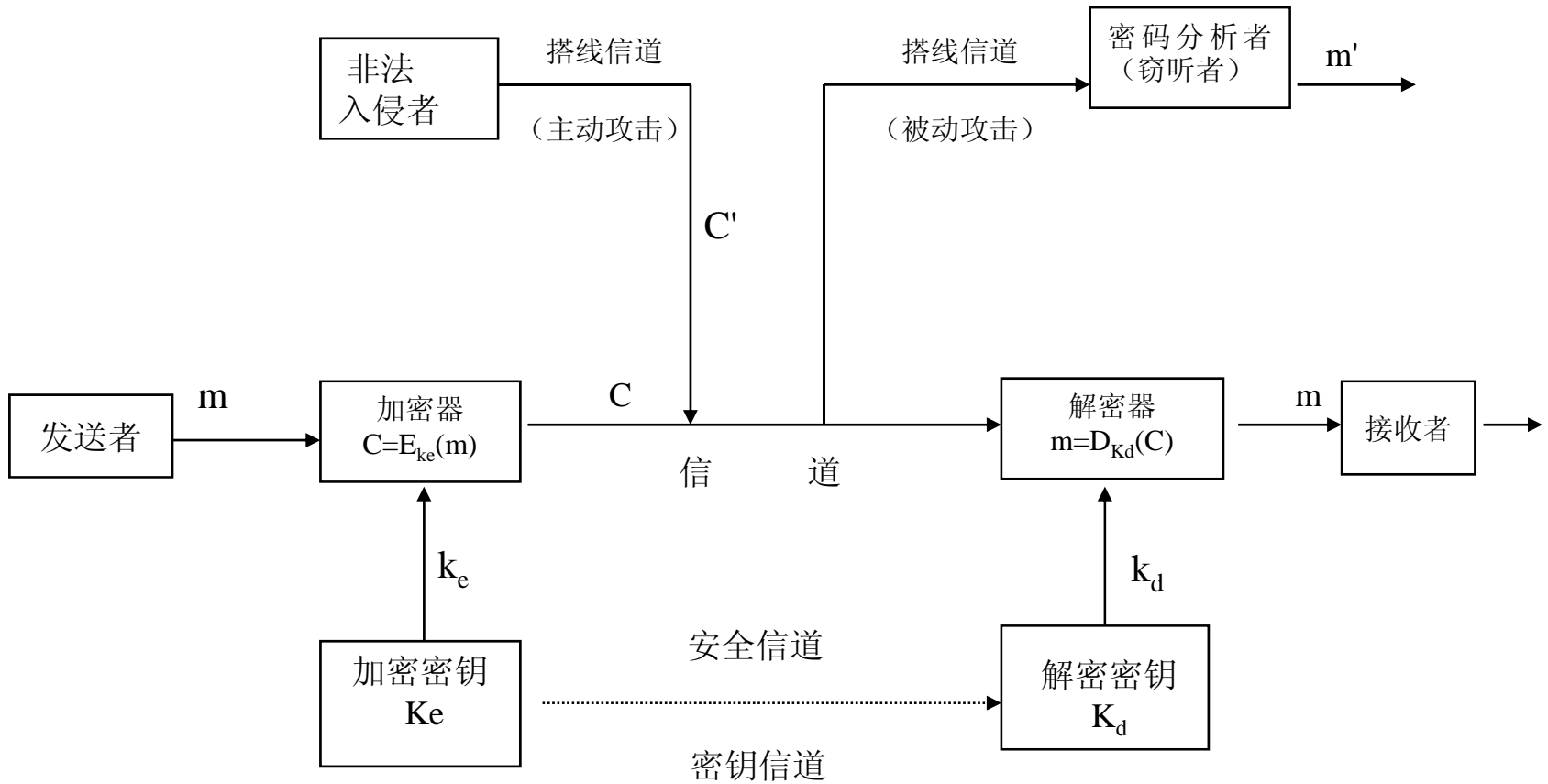


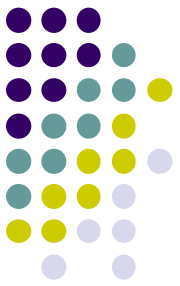
$$E_{K1}(M)=C, \quad D_{K2}(C)=M, \quad D_{K2}(E_{K1}(M))=M$$

- 对称密码体制（私钥密码体制）：**K1=K2**
- 非对称密码体制（公钥密码体制）：**K1≠K2**



一般保密系统模型





替换密码

- 凯撒 (Caesar) 密码

明文: **Caesar cipher is a shift substitution**

密文: **FDHVDU FLSKHU LV D VKLIW VXEVWLWXWLRQ**

- 维吉尼亚 ([Vigenere](#)) 密码

明文: **The mark of the immature man is that he wants to die nobly for a cause while the mark of the mature man is that he wants to live humbly for one**

参数: **cipherkey**

密文: **vpt teiu sd vpt pqdkxstm bhr zc xfc b wl arxxq vw spi eyfja ndy e tky qg ewppv dlc oigr sw dlc oiibvv wel ka ioek ri ucviz xf vmtg pjtfci jmt wcl**



换位密码

密文：卑者高者鄙的尚的是通是墓卑行高志鄙证尚铭

明文：



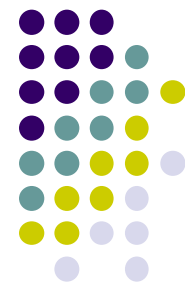
换位密码

密文：卑者高者鄙的尚的是通是墓卑行高志鄙证尚铭

明文：

卑	鄙	是	卑	鄙
者	的	通	行	证
高	尚	是	高	尚
者	的	墓	志	铭

一次世界大战时德国人用的ADFGVX密码就是换位密码与简单的替换密码的组合



密码系统的安全性

- 要说清楚安全性的含义，必须先定义下面两件事情：
 - 攻击者对密码系统的访问能力到底有多大，可以知道那些东西，不能知道哪些东要；
 - 攻击者要做到什么程度才算一个密码系统被攻破。
- 1883年, Kerckhoff指出：一个好的加密系统只能依赖密钥的秘密性，而不能依赖算法的秘密性。这在今天仍然是一条重要的原则。



什么是安全？

- 20世纪40年代, **Shannon**首次给出了完善保密的概念并证明了要达到完善保密, 密钥长度不能小于加密消息的长度。
- 1976年, **Diffie**和**Hellman**建议将密码系统的安全性建立在计算困难性问题的不好求解的性质之上, 第一次提出设计不容易破译的系统, 而不是设计不可能破译的系统。
- 在20世纪80年代, **Goldwasser**和**Micali**给出了第一个严格的令人满意的安全定义, 把密码的安全性严格建立在了一些广泛接受的计算假设之上。



什么是信息？

- 信息与不确定性
- 信息熵

$$H(X) = -\sum_{i=1}^n p(X = x_i) \log_2 p(X = x_i)$$

- 条件熵、联合熵：

$$H(X | Y), \quad H(X, Y)$$



完善保密性

- 20世纪40年代，仙农 (Shannon) 在他的“保密系统的信息理论”一文中，首次给出了完善保密性的概念。
 - ✓ 密码体制：(P, C, K, E, D)
 - ✓ 完善保密： $H(P|C)=H(P)$
- Gilbert Vernam “一次一密”密码体制(1917)
- 完善保密：密钥长度不能小于加密消息的长度
- 伪密钥与唯一解距离



启发与思考

- 保密性与随机性
- 包含信息就一定能获取到吗？
- 计算信息论(A. Yao, 1982)
- 单向函数
- 陷门单向函数



计算安全性

- 计算安全性始于1976年Diffie和Hellman发表的“密码学的新方向”一文，他们建议将密码系统的安全性建立在计算上的难问题的不好处理的性质之上；
- 1978, Rivest, Shamir 和Adelman根据Diffie和Hellman设计思想，首次提出了一个具体的公钥加密体制RSA。
- Diffie-Hellman的工作也可以看作是信息论思想的一种发展,导致了计算信息论的提出.



可证明安全性

- 在20世纪80年代，Goldwasser和Micali给出了第一个严格的令人比较满意的安全性定义，类似于从公理出发，他们把密码的安全性严格建立在一些广泛接受的计算困难性假设之上。
- 可证明安全的基础是计算复杂性理论，核心是 $NP \neq P$ ，严格说是 $NP \setminus BPP \neq \emptyset$ 。
- 由此奠定了现在密码学的基础



现代密码学

- 单向函数
- 伪随机性(伪随机函数、伪随机置换)
- 计算不可区分性
- 安全性证明
- 密码学可信问题
- 零知识证明系统



单向函数

- **可忽略的函数 $u(n)$** : 对任意的多项式, 对于充分大的 n , 都有

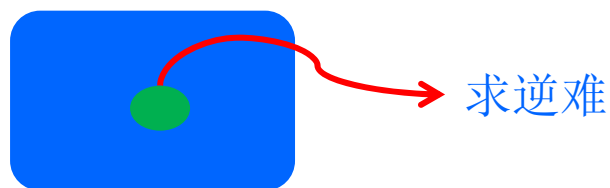
$$|u(n)| < 1/p(n)$$

- **单向函数 $f(x)$** : 存在一个多项式算法 A , 使得 $A(x) = f(x)$, 但对任意的概率多项式算法 D , $Pr[D(f(U_n), 1^n) \in f^{-1}(f(U_n))]$ 是一个可忽略的 n 的函数。

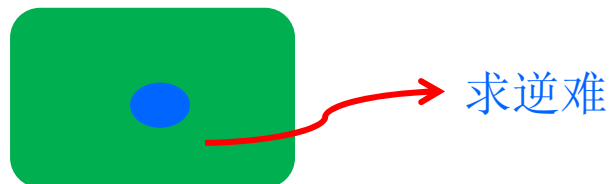


困难性与单向性

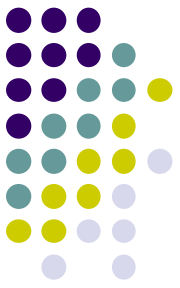
弱单向函数:



强单向函数:



- 问题: 怎样从弱单向函数构造强单向函数?



Yao方法

- Hardness Amplification:

$$F(x_1, x_2, \dots, x_t) = (f(x_1), f(x_2), \dots, f(x_t))$$

其中 $t = n \cdot \text{poly}(n)$

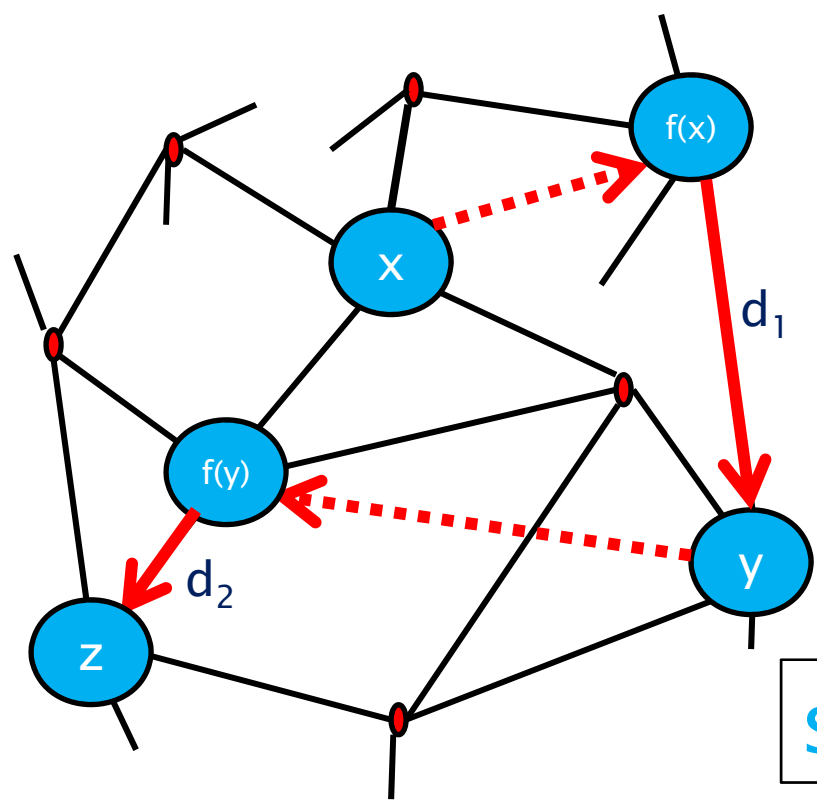
- Not security preserving

- 问题：给定一个弱单向函数，构造一个保安的强单向函数，即 $t = c \cdot n$ 对某一常数 c .



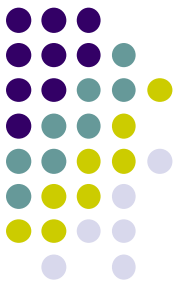
[GILVZ FOCS 90]

从单向置换 $f:\{0,1\}^n \rightarrow \{0,1\}^n$ 构造意向函数:



$$F(x, d_1, d_2, \dots, d_k) = (d_1, d_2, \dots, d_k, \bullet)$$

security-preserving!



伪随机性

- **计算不可区分性:** 两个概率分布 $X=\{X_n\}_{n \in N}$ 和 $Y=\{Y_n\}_{n \in N}$ 称为计算不可区分的, 如果对任意的概率多项式算法 D 、任意的多项式 $p(\cdot)$ 和足够大的 $n \in N$

$$|\Pr[D(X_n, n)=1]-\Pr[D(Y_n, n)=1]| < 1/p(n)$$

- **伪随机性:** 一个概率分布 $X=\{X_n\}_{n \in N}$ 称为伪随机的, 如果存在一个一致分布 $U=\{U_n\}_{n \in N}$ 使得 X 和 U 是多项式时间不可区分的。



困难性与随机性

- 伪随机数发生器：
是一个确定性的算法，它将一个短的随机消息（称为种子）扩展成一个长的看起来随机的序列。
- 定理（HILL 99）：
我们可以从任何一个单向函数构造一个伪随机数发生器



困难性、单向性与伪随机性

单向函数存在 \iff 伪随机数生成器的存在

单向函数的存在依赖于 $NP \setminus BPP \neq \{\}$ 假设

公开问题： $NP \setminus BPP \neq \{\}$ 能推出单向函数存在吗？



攻击模型

- 在现代密码学中，我们一般总是假定攻击者可以询问解密预言机或加密预言机，得到一些密文和对应的明文或明文对应的密文。
- 考虑一个攻击者和一个加密系统在玩一个游戏：加密系统随机地选择两个明文 m_0 和 m_1 ，加密其中一个，并把加密结果和 m_0 、 m_1 一起发给攻击者。如果攻击者不能判断出加密的是哪个明文，加密系统赢，这时我们称该加密系统是**消息不可区分的**



安全性证明

所谓密码算法的安全性证明就是判定在上述攻击模型中，密码算法和所依赖的困难性问题之间的规约关系，如果算法的攻破(即攻击者赢)意味着某一在密码学中可信的困难性问题的解决，则说明该算法是安全的，即安全性得到证明。



可信密码学问题

- 基于因数分解的问题
 - 整数分解问题
 - RSA问题: 求 x , 使得 $y=x^e$
 - E次方根问题: 给定一个 e , 求 x , 使得 $y=x^e$
- 基于离散对数的问题
 - 离散对数问题
 - 计算性Diffie-Hellman问题: 计算 $\text{Log}_H = \text{Log}_{H1} \text{Log}_{H2}$
 - 判定性Diffie-Hellman问题: 判定 $\text{Log}_H = \text{Log}_{H1} \text{Log}_{H2}$
- 基于多变元代数方程求解的问题
 - MQ问题, 多项式同构问题



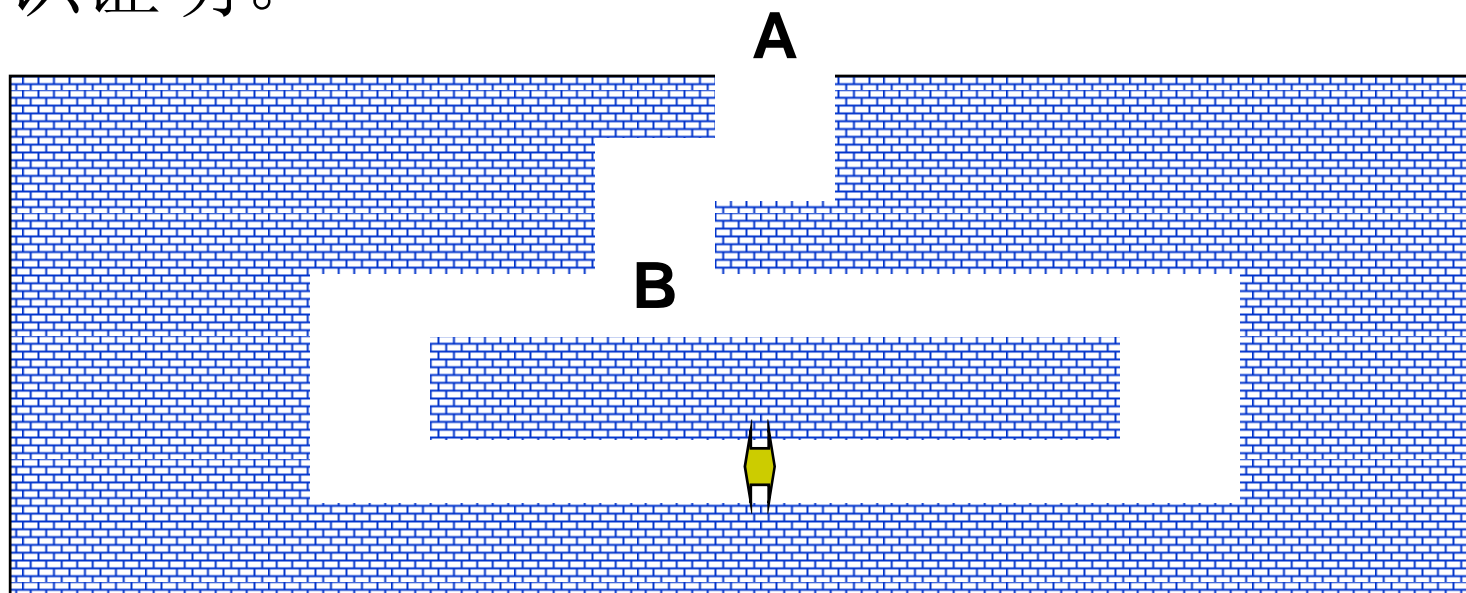
可信密码学问题

- Lattice约化问题
 - 最短向量问题
 - 最近向量问题
- 纠错码译码问题
- 问题：寻找其他密码学可信问题？
 - 平均复杂性
 - 陷门的隐藏
 - 空间复杂性的利用？



零知识证明

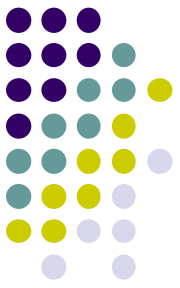
- 当事人的双方P和V，在协议执行后，如果V除了知道P能够证明某一事实外，没有得到任何关于这一事实的知识，我们则称P实现了零知识证明。





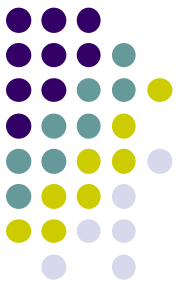
交互证明系统

- 一个集合 S 的交互证明系统是一个运行概率多项式时间算法的验证者 V 和具有无限运算能力的证明者 P 之间的一个两方协议，满足：
 - 完备性：对 S 中的任一元素 x ，与 P 交互之后 V 总能接受；
 - 合理性：对不属于 S 的 x ，无论采用什么样的证明策略 P^* ，与 P 交互后， V 总有至少 $1/2$ 的概率拒绝。



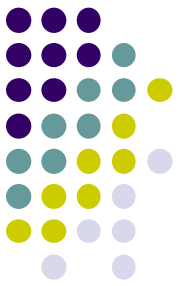
GNI的交互证明系统

- 公共输入: $G1 = (V1, E1)$, $G2 = (V2, E2)$
- $V1$: 随机选取一个图 Gx 和一个图的置换 S , 将 $S(Gx)$ 发给证明者 P ;
- P 收到 $G'=(V', E')$ 后确定 x 后并发送给 V ;
- $V2$: V 收到 x' 并验证与 x 是否相等, 相等则输出 1, 否则输出 0.



IP的威力

- $IP = PSPACE$
- 随机性+交互
- 公开掷币证明系统：验证者每轮的消息都是验证者在当前轮通过掷币的结果。
- 定理：任何一个有交互证明系统的集合也一定有一个公开掷币的证明系统



IP的层次结构

- IP(m)表示至多交换 $m(|x|)$ 个消息的语言类。
 $IP = IP(\text{poly})$, $NP < IP(1)$
- 线性加速: $IP(O(f(.))) \rightarrow IP(f(.))$, $IP(O(1)) \rightarrow IP(2)$, 其中 $f(.) \geq 2$.
- IP(2) contain GNI which not known to be in NP. $IP(2) = NP$ under plausible assumption.
- If $\text{coNP} < IP(2)$, then PTH collapse.
- $IP(f) < AM(f+3)$, $AM(2f) < AM(f)$



零知识证明系统

- 一个语言L的零知识证明实际上是一个交互式证明系统(P,V)，满足：

➤ **完备性**：任给 $x \in L$, $\Pr((P,V)(x)=1) > 2/3$

➤ **合理性**：对任意的证明者 P^* 和 $x \notin L$,

$$\Pr((P^*,V)(x)=1) < 1/3$$

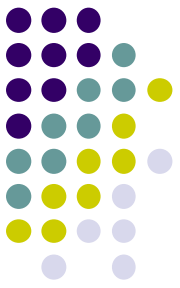
➤ **零知识性**：对任意的概率多项式时间验证者 V^* ，存在模拟器M，使得

$$\{\langle P, V^* \rangle(x)\}_{x \in L} = \{M(x)\}_{x \in L}$$



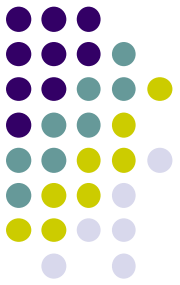
零知识系统的分类

- 根据证明者的能力区分：
 - 证明者的能力是无限的：零知识证明系统；
 - 证明者的能力是多项式时间的：零知识论证系统；
- 根据模拟的程度区分：
 - 模拟和真实交互是计算不可区分的：计算零知识协议；
 - 模拟和真实交互是统计不可区分的：统计零知识协议；
 - 模拟和真实交互是一样的：完美零知识协议；
- 非交互零知识、黑盒模拟、非黑盒模拟



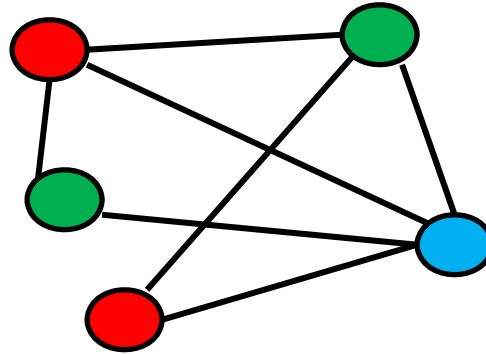
交互证明与零知识系统

- 两个概念提出都产生的具大的影响
 - 计算机科学
 - 密码学
- 自提出以来
 - 14 (among 42) Godel medalists for their work related to this field
 - 4 (among 22) ACM Doctoral Dissertation Award for their theses related to this field



零知识证明系统

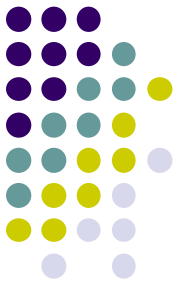
An NP-complete problem: Graph-3-Coloring



Graph 2-coloring : *easy*

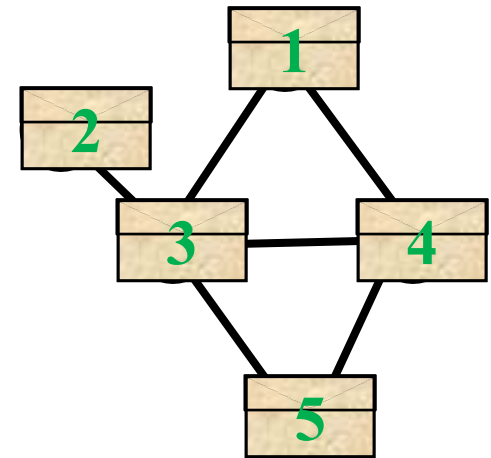
Graph 4-coloring : *trivial*

G3C零知识证明系统

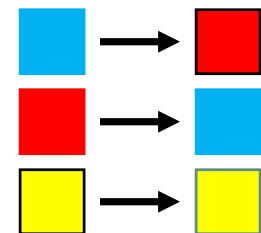


0. Prover chooses a random color permutation

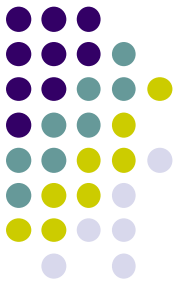
1. Prover puts all the vertices colors inside envelopes And sends them to the verifier.



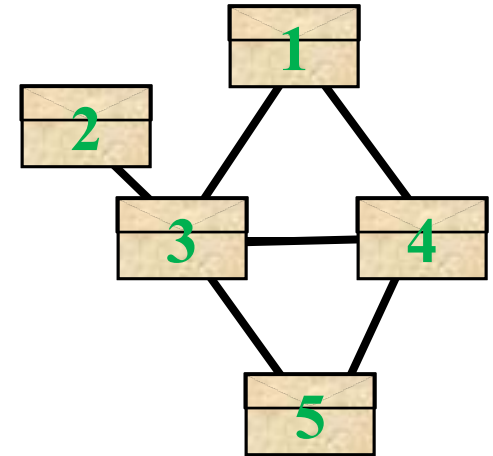
2. Verifier sends a query edge, say (4,5).



G3C零知识证明系统



3. Prover opens the envelopes, revealing the colors.
4. Verifier accepts if the colors are different.





ZK的威力

- **定理：**假定单向函数存在的情况下，对任意的 $L \in \text{NP}$ ，都存在关于L的零知识证明系统。
- **定理：**假定单向函数存在的情况下， $\text{IP} = \text{ZK}$
- **问题：**其他的性质、效率、安全性



零知识协议的通信复杂性

- 零知识协议的复杂性：
 - 轮数：常数轮？
 - 消息交换的总量
- 图三着色语言的零知识协议不是常数轮的：这是因为为了满足合理性， $O(n)$ 轮是必须的（是公共输入的长度）；
- 那么，是否存在常数轮的零知识协议呢？



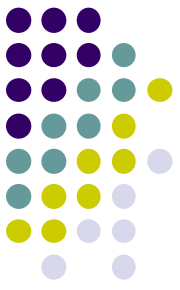
常数轮的零知识协议

- Brassard, Crepeau和Yung在论文[BCY89]中第一次给出了常数轮的零知识**论证**系统;
- Fiege和Shamir在论文[FS90]中为NP构造了四轮的零知识交互**论证**系统;
- Goldreich和Kahan在论文[GKa95, GKa96]中为NP构造了五轮的零知识**证明**系统, 作者采用了两个承诺方案, 一个用于限制验证者“作弊”, 而另一个则是用于限制证明者“作弊”的;



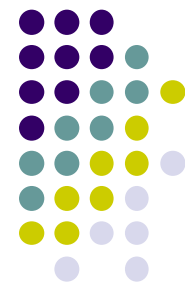
常数轮的零知识协议

- 是否存在四轮、三轮甚至两轮的零知识协议呢？让我们从反方向来说明这个问题：
 - Goldreich和Oren在论文[G094]中证明了，只有平凡的语言才拥有一轮和两轮的零知识协议；
 - 其次，Goldreich和Krawczyk在论文[GK96]中证明了，只有平凡的语言才拥有三轮的、公有掷币的、黑盒模拟零知识协议，并且证明了只有平凡的语言才拥有常数轮的、私有掷币的、黑盒模拟零知识协议（包括证明系统和论证系统）；
 - **问题：**构造三轮的零知识协议：实际上，根据[GK96]的结果，问题等价于，构造三轮的非黑盒零知识协议；



零知识系统的合成安全性

- 零知识系统的合成：
 - 顺序合成；
 - 并行合成；
 - 并发合成；
- Goldreich和Krawczyk在论文[GKr96]中证明了计算零知识协议在顺序合成下是封闭的
- 论文[FS90, Gkr96]分别给出了计算零知识协议并行合成并不封闭的反例；
- **问题：其他零知识系统的合成安全性？**



并发合成安全性

- [KPR98]证明了只有BPP中的语言才拥有四轮的、并发的、黑盒模拟的计算零知识协议，因此给出了一个下界；
- [DL08] 在**BPK** 模型中构造了对所有**NP**语言的常数轮并发合理的可重置零知识证明系统，并且只假设存在抗多项式时间敌手攻击的**hash**函数。
- **问题**: 是否存在对NP语言的常数轮的并发零知识论证系统？
- **问题**: 是否存在公开掷币的并发零知识论证系统？



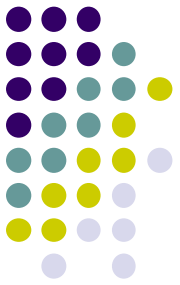
双重可重置猜想 (BGGL)

- **猜想：**是否存在对**NP**语言的重置合理的可重置零知识论证系统？
- [DL07] 通过引进实例依赖的可验证随机函数，构造了第一个常数轮具有某种特定的双重可重置性的零知识证明系统(弱重置合理，类-有界可重置零知识，可以是重置次数无界)，这里假设存在抵抗多项式时间敌手攻击的hash函数和陷门置换。给出了双重可重置猜想成立的一个充分条件，即证明了证明了公开掷币的并发零知识隐含了双重可重置零知识。
- [DGS09] 通过改进上述工具，彻底证明了双重可重置猜想



密码学的层次结构

- 第一层：由计算难问题组成，它们是密码学基本构建石块。一般来说，这需要用到还没有被证明的 $NP \neq P$ 的假设。
- 上述“难”问题被用来装配密码学基本构件，即执行一些基本任务，如加密、签名、随机数生成器、单向函数和Hash函数等的数学函数。
- 这些密码学基本构件将被用来组装密码协议，如通信、签名、拍卖、选举、电子支付、零知识证明等。这些仍然在安全的严格数学定义和证明的范围之内。
- 第四层是安全系统。安全系统实现安全协议，需要考虑现实生活中的限制以及数学模型之外的攻击。



谢谢大家！