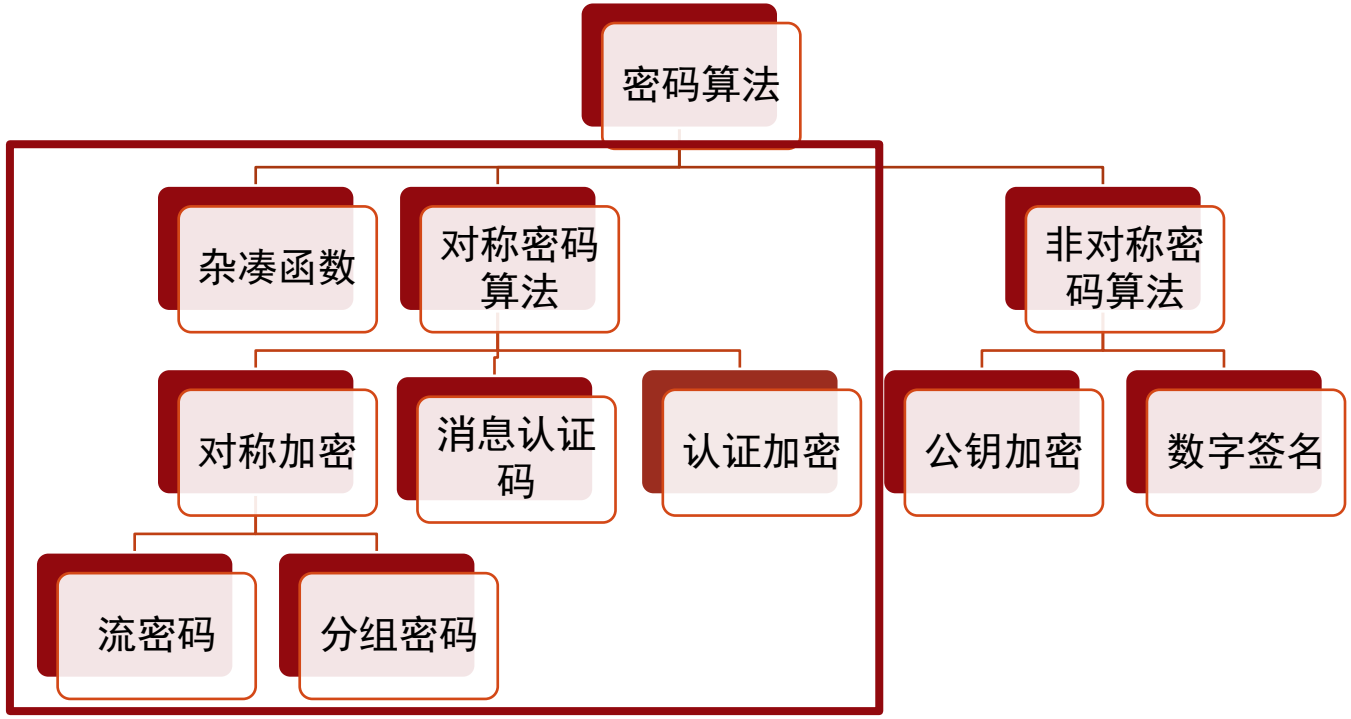
第1章密码分析学概述

密码学

- 密码学包括密码设计学和密码分析学
  - 密码设计者致力于设计出安全高效的密码算法（防），保护明文和密钥。
  - 密码分析者力图找到算法的某些安全缺陷（攻），尝试打破设计者宣称的安全界限，恢复明文或者密钥信息。



密码分析学

- 密码设计和密码分析是共生的、又是互逆的，两者密切相关但追求的目标相反。
- 两者解决问题的途径有很大差别。密码设计是利用数学来构造密码；密码分析除了依靠数学、工程背景、语言学等知识外，还要靠经验、统计、测试、眼力、直觉判断能力，有时还靠点运气。

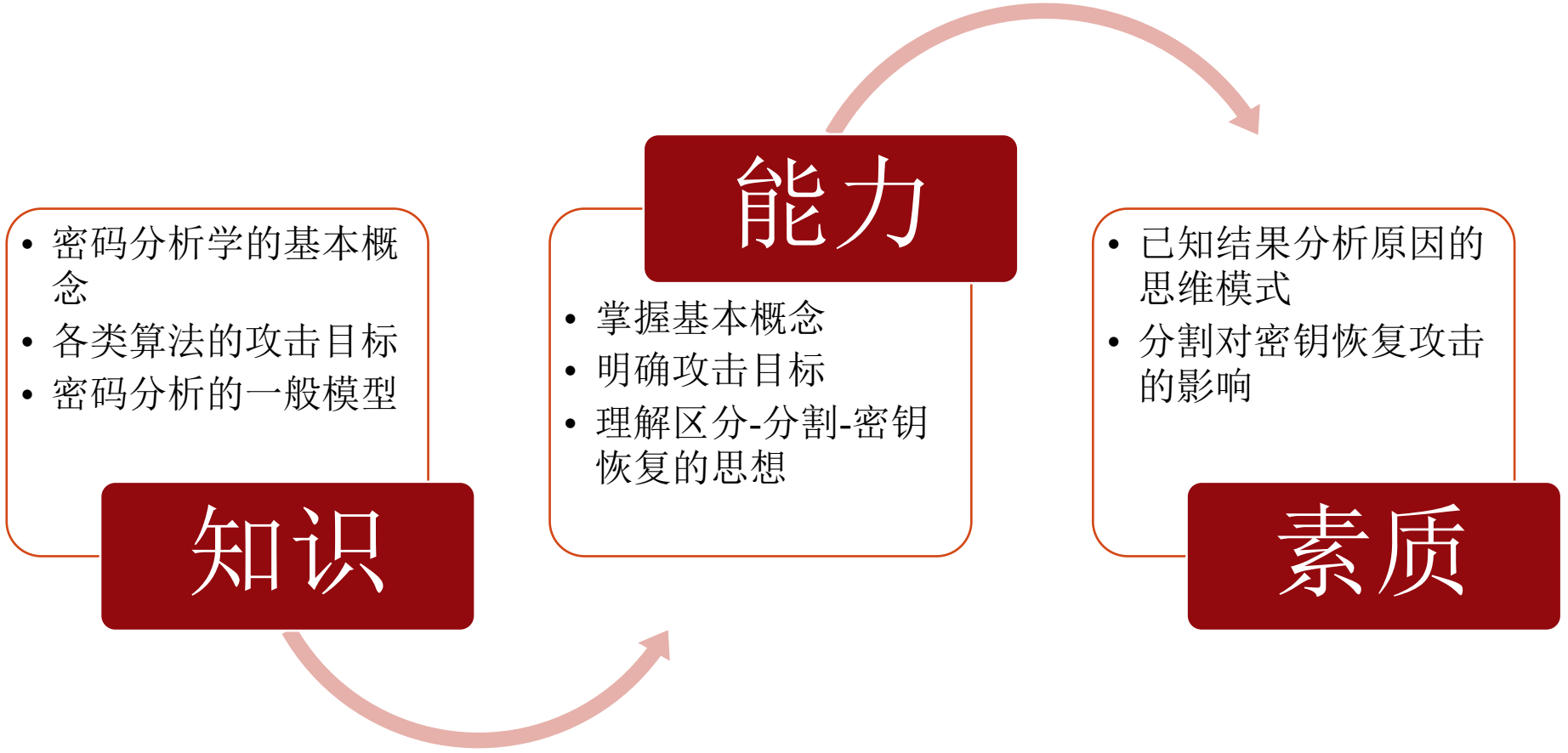
密码分析学


密码分析学
密码分析学概述
基本概念、算法攻击目标、密码分析的一般模型
古典密码破译
单表代换密码：1. 移位代换密码 2. 乘数密码 3.仿射密码 4. 多项式代换密码 5. 密钥短语密码
多表代换密码：1. 多字母代换密码 2. 维吉尼亚密码
恩尼格玛机破解
1. 猜测明密文的对应、2. 恢复扰频器的设置 3. 恢复线路接线板的设置 4. 密钥恢复攻击
对称密码分析
1. 序列密码分析 2. 分组密码分析 3. 杂凑函数
公钥密码分析
整数分解:1. Fermat法 2. 连分数法 3. 筛法 4. Pollard法
离散对数：1. 大步小步法 2. Silver-Pohlig-Hellman算法 3. 指标法
可证明安全：1. 语义安全的公钥密码体制 2. 基于身份的密码体制 3. 基于属性的密码体制 4. 抗泄露的公钥密码体制
密码协议安全性分析
1. 秘密共享（信息与图像） 2. 不经意传输 3. 电子投票 4 . 零知识证明 5. 电子现金支付
应用密码安全性分析
基于密码学的访问控制协议，包括基于密钥管理、公钥广播加密、属性加密的访问控制
云存储安全检索协议，包括对称密文检索、非对称密文检索、密文区间检索
安全协议：同态加密技术、可验证计算技术、安全多方计算技术、函数加密技术、外包计算技术
隐私保护协议：关系型数据隐私保护、社交图谱中的私保护、位置轨迹隐私保护、差分隐私

课程考核

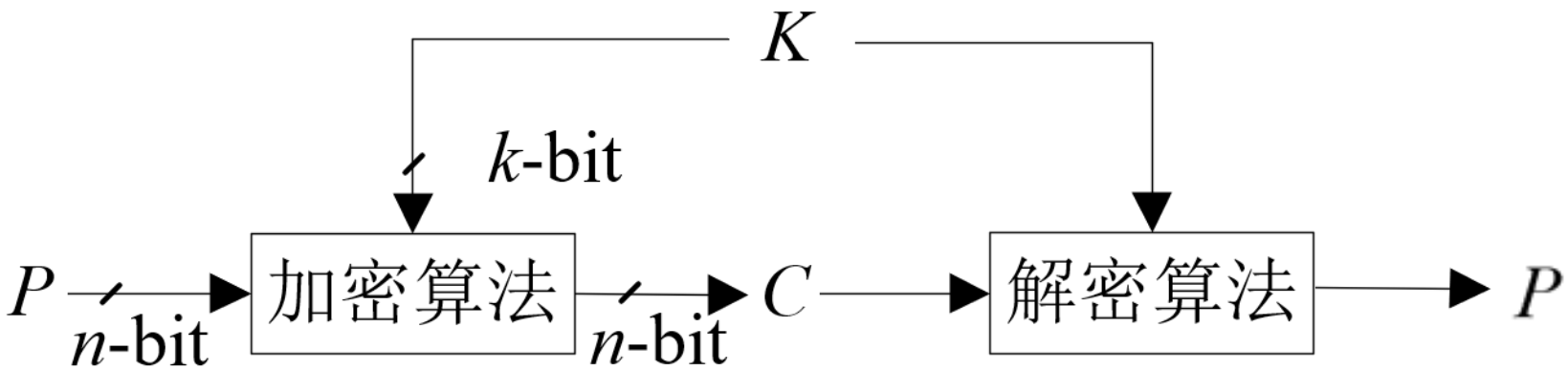
- 平时成绩
  - 10%
  - 签到&随堂测验（随机课堂点名缺勤扣2分）
- 实验成绩
  - 50%
  - 3个课后作业和1个课堂报告（第五周和第六周，每人10-12分钟）
- 期末成绩
  - 40%
  - 报告考察

教学目标


知识
密码分析学的基本概念
各类算法的攻击目标
密码分析的一般模型
能力
掌握基本概念
明确攻击目标
理解区分-分割-密钥恢复的思想
素质
已知结果分析原因的思维模式
分割对密钥恢复攻击的影响

对称加密算法（流密码，分组密码）

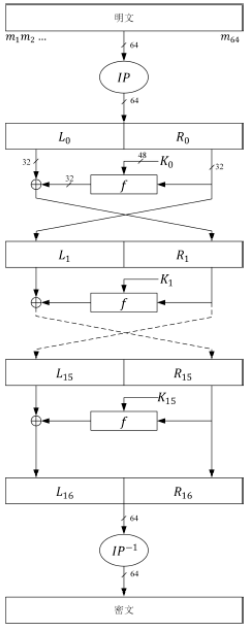
- 消息收发双方共享密钥

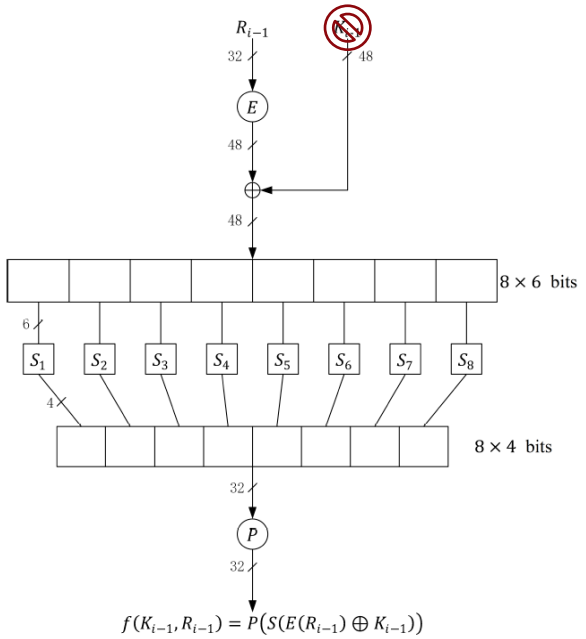


第1章密码分析学概述

基本假设

- 对称加密算法
- Kerckhoffs准则（Kerckhoffs'sprinciple）
  - 密码体制的安全性仅依赖于密钥，其他一切(包括算法本身)都是公开的
  - 密码算法的安全性应完全依赖于密钥的保密性，而非算法本身的保密性





攻击类型　　敌手能力

- 根据攻击环境的不同划分
  - 唯密文攻击(Ciphertext-onlyattack)：密码分析者能利用的资源仅为同一密钥加密的一个或多个密文，这是对密码分析者最不利的情况（**截获的部分密文**）
  - 已知明文攻击(Known-plaintextattack)：密码分析者者能够获得某些明密文的对应关系，这是密码算法至少需要抵抗的一种攻击（**截获的部分密文和对应的明文**）
  - 选择明文攻击(Chosen-plaintextattack)：密码分析者能够选择明文并获得相应的密文，这是对密码分析者十分有利的情况（**加密黑盒子，可加密任意明文得到相应的密文**）
  - 选择密文攻击(Chosen-ciphertextattack)：密码分析者能够选择密文并获得相应的明文，这也是对密码分析者十分有利的情况（**解密黑盒子，可解密任意密文得到相应的明文**）
  - 选择文本攻击（Chosentextattack）：密码分析者能够选择明文并获得相应的密文也能够选择密文并获得相应的明文（**加密黑盒子和解密黑盒子**）

已知明文攻击举例

要抵抗已知明文攻击，必须精心地设计加解密算法（E，D）。（能抵抗已知明文攻击的加解密算法(E,D)并不是很容易构造的。）

例1 设：加密密钥等于解密密钥：$z=k$；加密算法为$c= m+z$；对应的解密算法为$m=c-k=c-z$。（普通加减法）

注意到此时$k=c-m$。这就是说，只要知道了一组明文/密文对（m，c），就能计算出解密密钥k。

已知明文攻击举例

例2 设：加密密钥等于解密密钥，$z=(z_1, z_2)$；

加密算法为$c=z_1m+z_2$；对应的解密算法为$m=(c-z_2)/z_1$。（普通加减乘除法）

设攻击者Eve获得了以往废弃的2组明文/密文对：（m_1，c_1），（m_2，c_2）。注意到此时

$$c_1=z_1m_1+z_2;$$

$$c_2=z_1m_2+z_2。$$

这是一个关于密钥(z_1, z_2)的 二元一次方程组，能计算出(z_1, z_2)。

无条件安全和计算安全

无条件安全（完美保密）

对密码体制的任何攻击，都不优于（对明文）完全盲目的猜测，这样的密码体制就称为无条件安全的（或完善保密的）。

一次一密的加密方式容易实现无条件安全性。因为密钥时时更新，所以以往得到的任何明文/密文对，对于破译新的密文没有任何帮助，只能做完全盲目的猜测。

无条件安全和计算安全

计算安全

计算安全是一个模糊的概念。我们可以给出以下三个级别的定义。

（1）对密码体制的任何攻击，虽然可能优于完全盲目的猜测，但超出了攻击者的计算能力。这是最高级别的计算安全。

（2）对密码体制的任何攻击，虽然可能没有超出攻击者的计算能力，但所付出的代价远远大于破译成功所得到的利益。这是第二级别的计算安全。

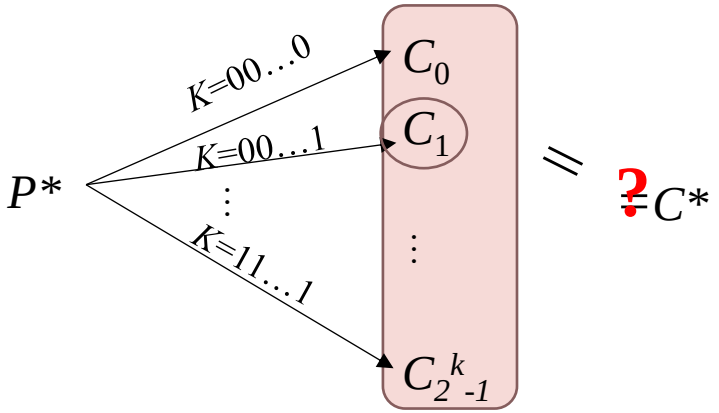
（3）对密码体制的任何攻击，虽然可能没有超出攻击者的计算能力，但破译成功所需要的时间远远大于明文本身的有效期限。这也是第二级别的计算安全。

通用的攻击——穷举攻击

- 对截获的密文，依次用各种可能的密钥试译，直到得到有意义的明文。各种可能的密钥的总数称为密钥量。
- 只要有足够多的计算时间和存储容量，原则上穷举搜索总是可以成功的。任何一种实用密码的密钥量都远远大于攻击者所能承受的计算时间和存储容量。

通用的攻击——穷举攻击

- 强力攻击的一种
- 敌手获得一个明密文对(P^*, C^*)
- 正确密钥的判定条件：$E_k(P^*) = C^*$



- 攻击类型？
- 已知明文攻击
- 是否实际可行？建方程、解方程需要的计算资源？

通用的攻击——穷举攻击举例

- 分组长度为64比特，密钥长度为80比特
- $\forall k, \text{s.t.}\ E_k(P^*) = C^*$的概率为$\frac{1}{2^{64}}$（随机假设下）
- 若穷举全部的密钥进行验证，则一个明密文期望筛选出
- $2^{80} \times \frac{1}{2^{64}} = 2^{16}$个候选密钥
- 如何筛选出唯一正确的密钥？
- 再获取一个明密文！
- 没被筛除的错误密钥期望有$2^{16} \times \frac{1}{2^{64}} = 2^{-48}$个
- 建方程、解方程需要的计算资源
  - 时间：2^{80}次加密；数据：已知2个明密文；存储：2个明密文
- 成功率：100%
- 若穷举一半的密钥（2^{79}）呢？

- 时间：2^{79}；数据：2；存储：2
- 成功率：50%

计算安全

- 成功率Ps：攻击目标达成的概率
- 攻击复杂度：
  - 数据D：实现攻击所需的明文或密文的总数（个）
  - 时间T：对采集到的数据进行分析和处理所消耗的时间（加密次数）
  - 存储M：实现攻击占用存储空间的大小（B）
- 一般情况下，比较两个攻击优劣时，应把复杂度统一在相同的成功率下，再进行比较
- 若在不可忽略的成功率下，各攻击方法的复杂度均超出了分析者的计算资源可达到的合理边界，则称该密码体制是计算安全的
- 通用攻击给出安全上界（攻击复杂度的上界）
- 所以密码分析的目的是什么？不断降低安全上界

第1章密码分析学概述

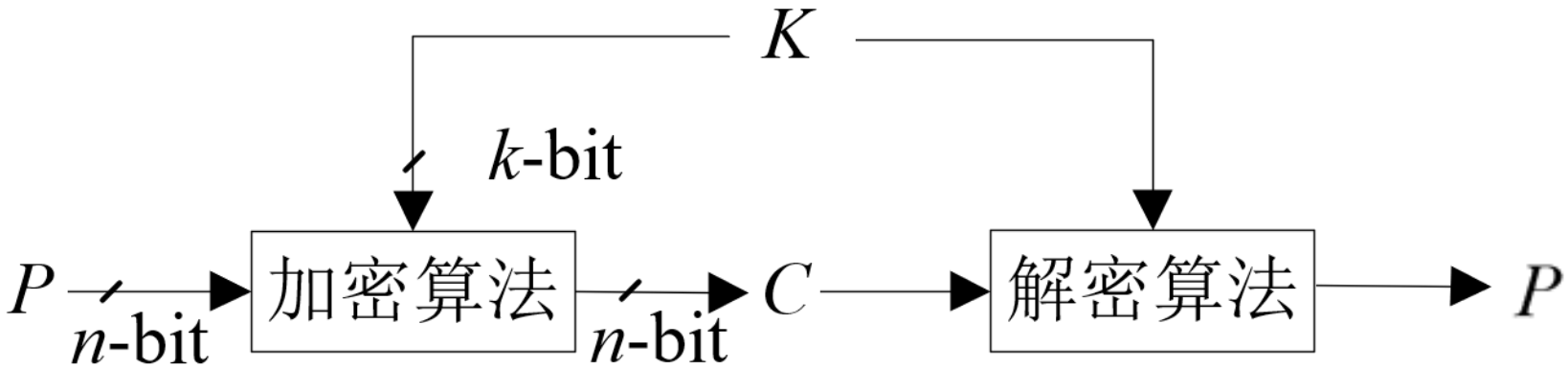
1.1密码分析学的基本概念

1.2各类算法的攻击目标

1.3密码分析的一般模型

对称加密算法（流密码，分组密码）

- 安全属性
  - 主要用于保障机密性，同时，也是随机数生成器、杂凑函数或消息认证码等算法的基本部件,因此,对其进行安全性分析，主要考以下两类攻击.
- 攻击目标
  - 区分攻击：将对称加密算法与随机置换进行区分
  - 密钥恢复攻击：恢复出分组密码算法进行加解密运算采用的密钥
- 复杂度上界2^k



杂凑函数（哈希函数或散列函数）

- 安全属性
  - 保障消息完整性，同时也是数字签名的关键部件，对杂凑值长度为n比特的杂凑算法h，针对其安全属性，主要考虑四种攻击。
- 攻击目标

(i)原像攻击:给定n比特的哈希值H,找到消息M,满足h(M)=H。

(ii)第二原像攻击:给定消息M_1，找到另一个数据串M_2，满足$h(M_1)=h(M_2)$且$M_1 \neq M_2$.

(iii)碰撞攻击:找到两个消息(M_1,M_2),满足$h(M_1)=h(M_1)$且$M_1 \neq M_2$。

(iv)长度扩展攻击:给定n比特的杂凑值h(M),其中M为未知的非空数据串,找到任意数据串N和n比特的H’，满足h(M||N)=H’。（**由于该攻击的存在,MD结构的杂凑函数用于构造某些消息认证码或认证加密算法时，易于进行伪造,因此美国国家标准与技术研究院（NIST）在第三代杂凑函数标准SHA-3的征集过程中,要求抵抗该攻击**）

由于强力攻击的存在,原像攻击、第二原像攻击和长度扩展攻击的复杂度上界为2^n,而对于碰撞攻击,存在生日攻击,故复杂度上界为$2^{n/2}$.而对不同结构的杂凑算法,各类攻击的复杂度上界可能进一步降低。

消息认证码（MAC）

- 安全属性
  - 主要用于保障认证性,因其也视收发双方共享的密钥为密码体制中唯一保密的信息,故对称加密算法的攻击目标也适用于MAC算法此外,还可以考虑伪造攻击.
- 攻击目标
  - 区分攻击:
    - R型区分攻击(Distinguishing-RAttack):将MAC算法与随机函数进行区分。
    - H型区分攻击(Distinguishing-HAttack):将基于具体密码元件(如杂凑函数SHA-1)构造的MAC算法与基于随机函数构造的MAC算法进行区分。

消息认证码

- 攻击目标
  - **伪造攻击：** 设MAC算法的输入为M,输出为t,则不知道密钥的攻击者，输出能通过验证的(M,t),即Verf(M,t)=1.具体来说,分为以下三种,
  - **(i)存在性伪造(Existential Forgery):**攻击者与MAC算法进行交互后,输出(M,t),满足Verf(M,t)=1且M在交互过程中没有被询问(query)过,例如,攻击者自适应选择若干消息$M_1, M_2, \cdots, M_s$,访问MAC算法并获得对应的正确的$t_1, t_1, \cdots, t_s$,然后输出(M,t),满足$M \neq \{M_1, M_2, \cdots., M_s\}$且Verf(M,t)=1.
  - **(ii)选择性伪造(Selective Forgery):**攻击者在与MAC算法进行交互之前,选定一个消息M.然后,根据交互获得的信息,输出(M,t),满足Verf(M,t)=1日M在交互过程中没有被询问(query)过.
  - **(iii)通用性伪造(Universal Forgery):**对在与MAC算法进行交互之前任意给定的消息M,攻击者均可根据交互获得的信息,输出(M,t),满足Verf(M,t)=1目M在交过程中没有被询问过。

消息认证码

- **密钥恢复攻击**:恢复出MAC算法采用的密钥。
- 设MAC值t的长度为n比特,密钥长度为k比特,则区分攻击和伪造攻击的复杂度上界为$\min(2^n, 2^k)$,密钥恢复攻击的复杂度上界为2^k.而对不同结构的MAC算法,各类攻击的复杂度上界可能进一步降低.例如,文献[3,4]中给出的基于杂凑函数的部分MAC算法的区分攻击的复杂度上界为$2^{l/2}$.其中,l为中间链接变量的比特长度.
- 对于认证加密算法(AE)、公钥加密和数字签名算法都可类似考虑攻击者的攻击目标.
- 值得注意的是，密码分析并不仅仅分析完整的算法，一般会从分析算法的简化版本入手.例如,数据加密标准DES规定要16轮加密后的结果才是密文,那么在分析时,可以假设12轮就出结果,即对缩减到只加密12轮的DES的简化版本进行分析.如果发现算法的简化版本,通过分析算法在设计中存在的问题，随着时间的推移和技术的进步,可能会完成整个算法的破解.

第1章密码分析学概述

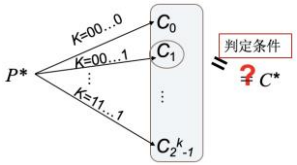
1.1密码分析学的基本概念

1.2各类算法的攻击目标

1.3密码分析的一般模型

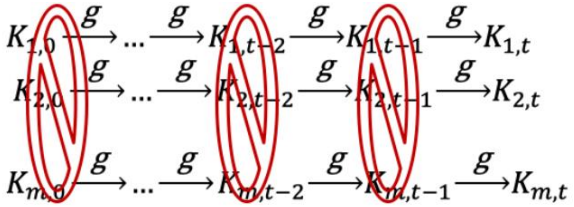
密码分析方法（多解）

- **强力攻击**：通用的攻击方法，与算法的设计细节无关
  - 给出算法的安全上界
  - 穷举攻击、字典攻击、查表攻击、时间-存储权衡攻击等强力攻击



明文	密文
M_0	C_0
M_1	C_1
…	…
M_{2^k-1}	C_{2^k-1}

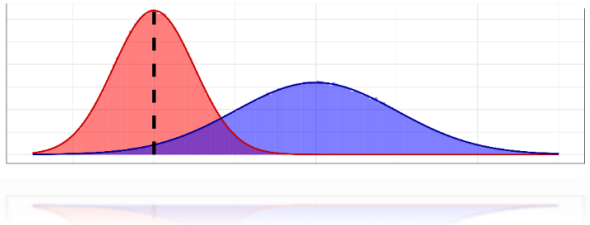
密文	密钥
C_0	00…00
C_1	00…01
…	
C_{2^k-1}	11…11



- **基于算法的设计细节的攻击方法**
  - 通过分析算法的内部结构和组件，将数学推导、统计测试与程序搜索相结合，发现特殊规律，开展分析工作
  - 差分分析、线性分析、积分分析等
- **基于算法实现时产生的物理参量的攻击方法（侧信道攻击）**
  - 算法在芯片中运行泄露的某些物理参量，如执行时间、电流、电压、电磁辐射、声音等信息与密码算法的中间状态数据和运算操作存在一定的相关性，攻击者通过采集这些泄露信息，推测密钥
  - 计时攻击、能量分析、电磁攻击等（时间开销、能量消耗、电磁信号强弱等不同）

攻击思想-区分

- 重点关注基于数学模型和程序搜索的攻击方法，关键要素是区分和分割。
- 区分
  - 不随机现象：一个可计算或统计的指标，该指标在具体的某个密码算法下的分布与在随机函数(或置换)下的分布不同，即在复杂度允许的范围内，能以不可忽略的概率区分这两种分布。



  - 区分攻击（区分器）：在复杂度允许的范围内，敌手利用不随机现象，以不可忽略的概率将密码算法和伪随机函数(或置换)进行区分的算法
  - 区分往往是开展各类攻击的第一步

区分攻击示例

- 某分组加密算法EG_2，P, C, K_0, K_1均为n比特，各密钥相互独立，F为置换
- 不随机现象：
  - 对加密算法：$P \oplus C = 1$
  - 对随机置换，$\Pr(P \oplus C = 1) = \frac{1}{2^n}$

- 如何进行区分攻击？成功率？
- 敌手需获取不随机现象相关的参数：P, C
- 已知明文攻击：已知x个输入P_i及相应的输出C_i

1. 计算x个$P_i \oplus C_i$；
2. 若x个值都等于1，则输出为EG_2算法；
3. 若x个值至少有一个不等于1，则输出为随机置换RP

区分攻击示例

- 成功率

1. 计算x个$P_i \oplus C_i$；
2. 若x个值都等于1，则输出为EG_2算法；
3. 若x个值至少有一个不等于1，则输出为随机置换RP

- 当黑盒的确为EG_2算法时，区分器输出为EG_2算法；当黑盒为随机置换时，输出为随机置换RP

- 假设$\Pr(\text{黑盒为}EG_2) = \Pr(\text{黑盒为RP}) = \frac{1}{2}$

$$\begin{aligned}
\Pr(\text{成功}) &= \Pr(\text{黑盒是 } EG_2 \text{ 且输出为 } EG_2) + \Pr(\text{黑盒是 RP 且输出为 RP}) \\
&= \frac{1}{2} \cdot \Pr(\text{输出为 } EG_2 | \text{ 黑盒是 } EG_2) + \frac{1}{2} \cdot \Pr(\text{输出为 RP} | \text{ 黑盒是 RP}) \\
&= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \left(1 - \left(\frac{1}{2^n}\right)^x\right) \\
&= 1 - \frac{1}{2^{nx+1}}.
\end{aligned}$$

$P(AB) = P(A)P(B|A)$

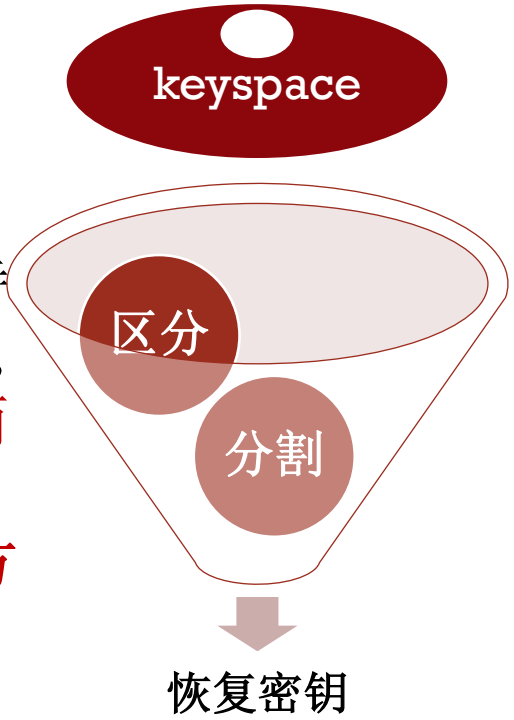
$x = 1$即可以不可忽略的概率区分成功

攻击思想——分割

- 密码算法的分割

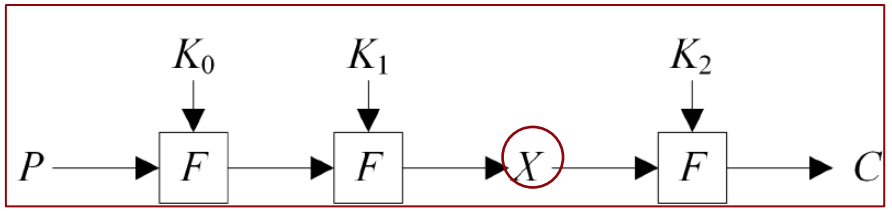
  指在寻找区分器时往往采用自下而上的研究思路，即将算法的各个部件分开考虑，细化到S盒、P置换，甚至细化到比特，来发现不随机现象。

- 密钥搜索空间的分割
  - 在进行强力攻击时，我们关注的是与全部密钥比特有关的明文和密文之间的关系，因此要验证该关系是否满足，必须考虑全部密钥的可能，那么，要改进攻击复杂度，就需要借助区分器，得到只与部分密钥比特有关的可验证的明文和密文之间的关系从而验证该关系是否满足时，只需考虑部分密钥，先对这部分密钥进行恢复，再恢复其余密钥，这种“两步走”的策略，对应的复杂度是“相加”的关系，从而降低总体的复杂度，以穷举攻击为例，**分割的目的是将密钥搜索空间进行“分割”，想方设法利用区分器降低搜索空间的大小或提高求解的效率，将必须整体搜索的大空间分割为可局部搜索的小空间.**
  - 将不随机现象转换为只与部分密钥比特有关的方程



密钥恢复攻击示例

- 某加密算法EG_3，P, C, K_0, K_1, K_2均为n比特，各密钥相互独立，F为置换

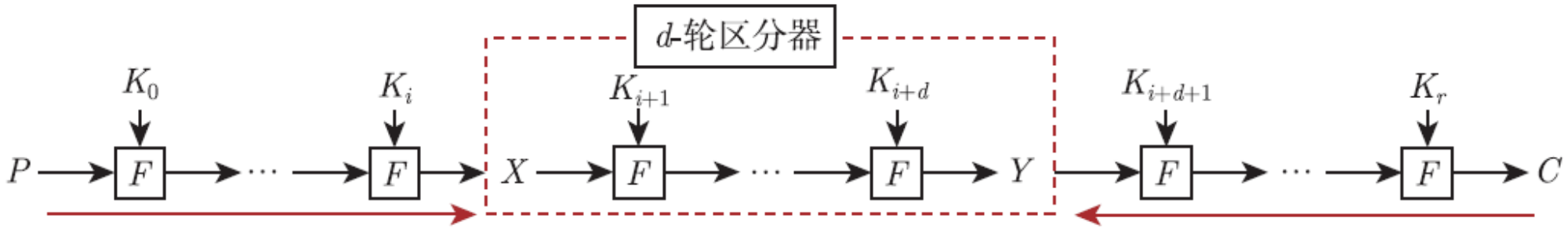


- 不随机现象（2轮的区分器）
  - $P \oplus X = 1 \Rightarrow X = P \oplus 1$
- $E_{K_0,K_1,K_2}(P^*) = C^* \Rightarrow$
- $C^* = F_{K_2}(X^*) = F_{K_2}(P^* \oplus 1)$
- 实现密钥空间的分割：先恢复K_2，再恢复K_0，K_1
- 复杂度$2^{3n} \rightarrow 2^{2n}(E_{K_0,K_1}(P^*) = X^*)$或者$2^n(E_{K_0}(P^*) = D_{K_1}(X^*))$

（穷举攻击或者中间相遇攻击）

密钥恢复攻击的一般模型（解题步骤）

- 第一步：找到一个d轮区分器
  - 区分器对应的不随机现象只与区分器的头尾(X,Y)相关，与每轮的中间值无关
  - 看做函数$D(X,Y)$
- 第二步：利用区分器实现密钥空间的分割
  - 由明文加密到区分器头部涉及的密钥$K_0,\dots,K_i$的部分比特，以及由密文解密到区分器尾部涉及的密钥$K_{i+d+1},\dots,K_r$的部分比特
- 攻击的复杂度及成功率与函数$D(X,Y)$及求解(X,Y)的方式和涉及的密钥量有关



其他攻击

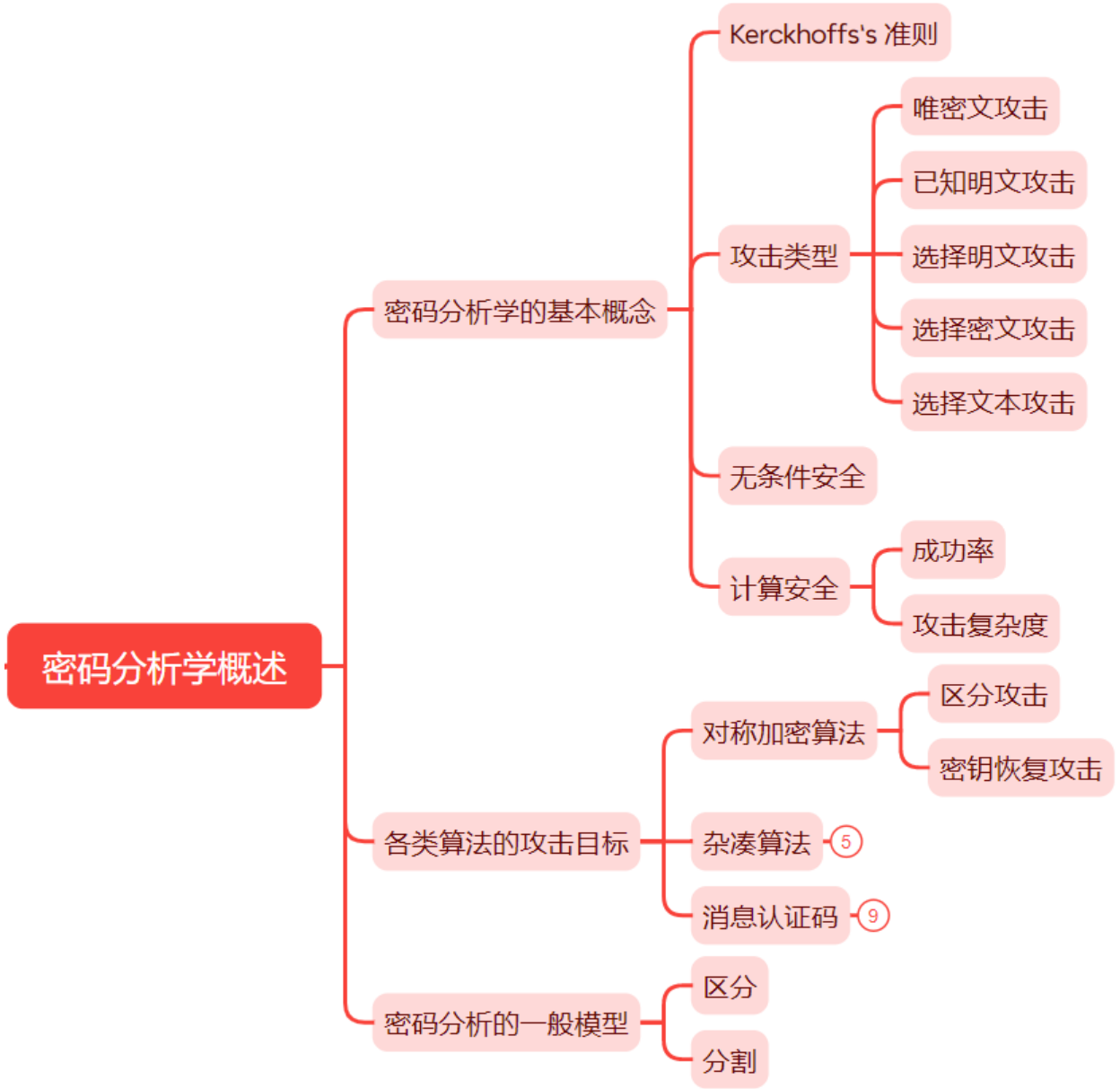
现代密码分析经过几十年的发展，出现了很多典型的分析技术，主要围绕区分器构造的不同来进行分类，

1. 差分分析利用高概率的差分路线构造区分器；
2. 在差分分析的启发下，飞去来器攻击和矩形攻击则同时利用两条高概率的差分路线构造区分器；
3. 截断差分分析进一步将取特定值的差分扩充为满足特殊条件的差分集合来发现不随机特性；
4. 不可能差分分析反其道而行之，根据概率为0的差分路线建立区分器，线性类分析技术则利用高偏差的线性逼近式构造区分器。

后来又出现了基于零偏差线性近似式的零相关攻击等，此外，还有中间相遇攻击、积分攻击等多种攻击方法，不同攻击方法的侧重点不同，对同一个算法的攻击效果不同。

衡量一个密码算法的安全性，往往需要考虑各种攻击方法的影响，特别是新的密码算法发布时，设计者必须给出利用现有各种攻击对算法进行安全性分析得到的详细数据，作为衡量算法安全性的重要指标.

小结


密码分析学概述
密码分析学的基本概念
Kerckhoffs's 准则
攻击类型
唯密文攻击
已知明文攻击
选择明文攻击
选择密文攻击
选择文本攻击
无条件安全
计算安全
成功率
攻击复杂度
各类算法的攻击目标
对称加密算法
区分攻击
密钥恢复攻击
杂凑算法
5
消息认证码
9
密码分析的一般模型
区分
分割

作业：预习恩尼格玛算法的破解

- 教材第2章及附录A.1
- 参考视频
  - 算法描述

    https://www.bilibili.com/video/av31393190/?spm_id_from=trigger_reload
  - 算法破解https://www.bilibili.com/video/av21919076/?p=2