2022-2023学年秋季学期

课程名称: 信息安全数学基础

英文名称: Mathematical Foundations for Information Security

授课团队：胡磊、许军、王丽萍

助　　教：郭一

信息安全数学基础

Mathematical Foundations for Information Security

[第 1 次课] 介绍

授课教师：胡磊

授课时间：2022年8月31日

概 要

- 课程介绍

信息安全在日常生活中

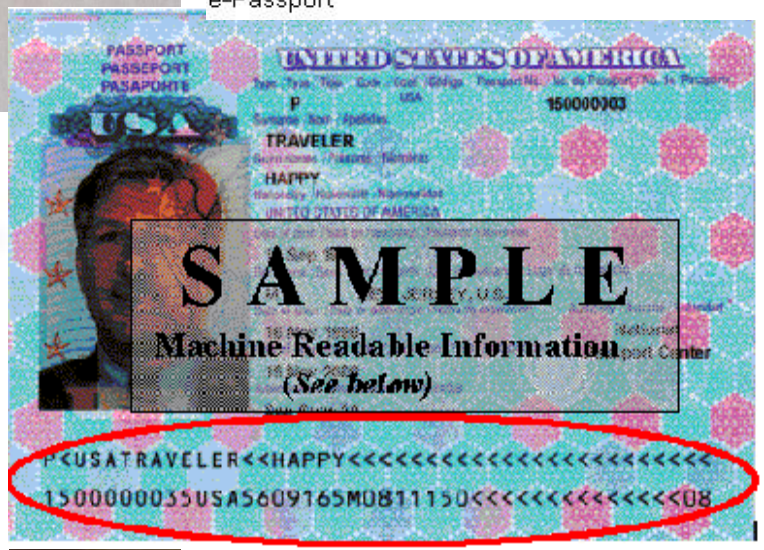
- 工行 U盾、农行二代K宝
  - 用于验证网银客户身份的安全工具，内置智能芯片采用复杂加密算法，确保银行客户的每一笔网银交易万无一失。尤其是对网上银行安全性有较高要求的个人客户，同时也适用经常性大额转账交易和网上支付的个人客户
- 电子护照
- https
- VPN
- 二代身份证


本护照内置敏感电子元件，
为保持最佳性能，请不要将护照
弯折、打孔或者暴露在极端温度、
湿度环境中。
This passport contains sensitive
electronics. For best performance,
please do not bend, perforate or
expose to extreme temperatures or
excess moisture.
请勿在此盖印
DO NOT STAMP HERE
中华人民共和国
People's Republic of China
护 照
PASSPORT


中 華 人 民 共 和 國 香 港 特 別 行 政 區
HONG KONG SPECIAL ADMINISTRATIVE REGION, PEOPLE'S REPUBLIC OF CHINA
PASSPORT
CHN
K12345599
鍾 / CHUNG
潤心 / KWOK SUM
CHINESE
F
08 AUG 80
HONG KONG
05 FEB 07
05 FEB 17
香港特別行政區入境事務處
IMMIGRATION DEPARTMENT, HONG KONG SPECIAL ADMINISTRATIVE REGION
P<CHNCHUNG<<KWOK<SUM<<<<<<<<<<<<<<<<<<<<<<<<
K123455994CHN8008080F1702057HK8888888<<<<<36

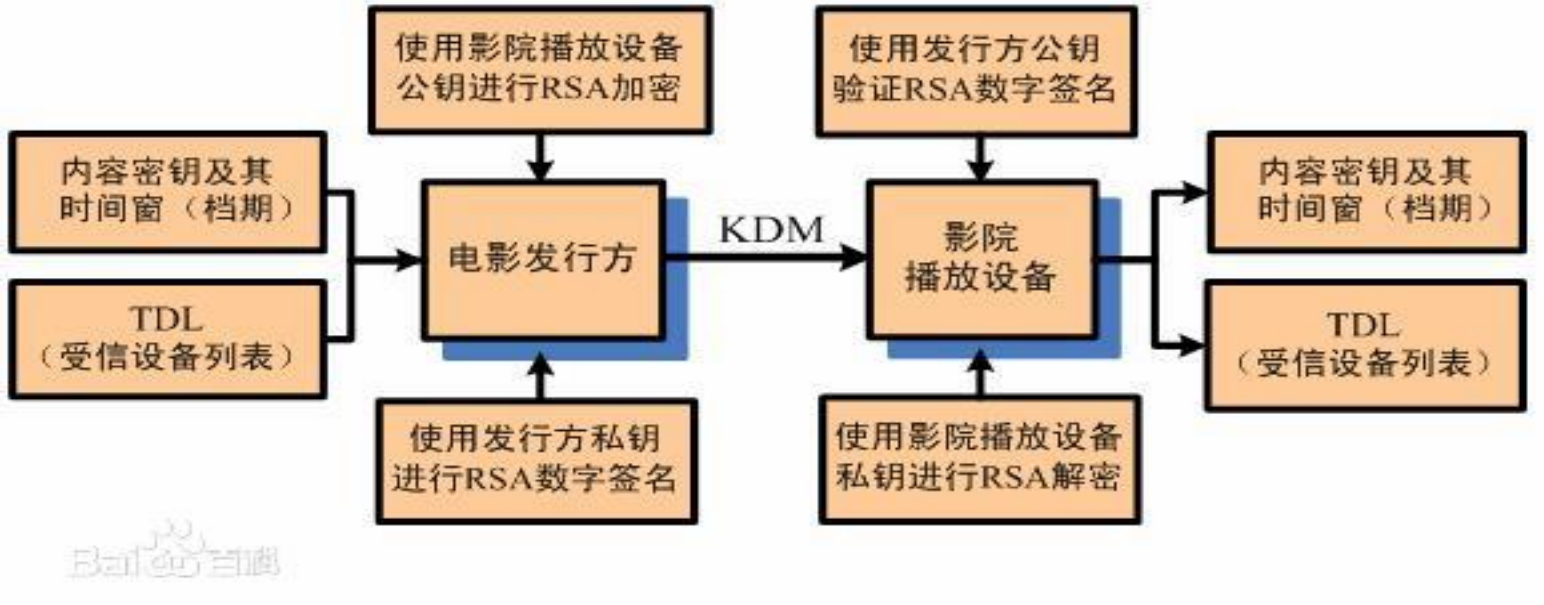
e-Passport


BUNDESDRUCKEREI
SPECIMEN ELECTRONIC PASSPORT


PASSPORT
PASSEPORT
PASAPORTE
USA
UNITED STATES OF AMERICA
P
USA
150000003
TRAVELER
HAPPY
SAMPLE
Machine Readable Information
(See below)
P<USATRAVELER<<HAPPY<<<<<<<<<<<<<<<<<<<<<<<<
150000003<USA5609165M0811150<<<<<<<<<<<<<<08


PASS
PASSEPORT
PASSAPORTO
PASSAPORT
护照

- 数字电影放映前，影院把数字电影拷贝即硬盘链接到服务器上，而密钥允许服务器读取影片。每一间影厅的服务器，都需要各自匹配的密钥，才能顺利放映
- 目前国内95%以上影片的密钥都是由中影数字电影发展有限公司制作。在电影上映前，该公司会将密钥放上官网的下载区，供影院下载。至于可以下载的时间，提前一周或半个月等不定。密钥有使用时间限制，一般有效期约为1个月，具体到每一部影片，都会有所不同



TLS1.3（2018/3/20最新版）

- 传输层安全性（ Transport Layer Security,TLS）是一种使网页浏览安全的协议
  - 网站向用户证实自己是真实的,不是钓鱼网站
  - 网站和用户建立连接后,必须保证传输数据的机密性和完整性：敌手无法获得安全信道上传输的信息， 敌手无法篡改数据,欺骗用户
- TLS1.3协议分为两个部分：
  - 一个是握手协议,进行实体认证,约定后续算法和使用的参数,建立共享密钥；
  - 另一个是记录协议(record protocol),对信道上传输的数据进行记录、分组、验证、解密、重组,然后把结果提交给更高层级的应用端.

TLS1.3使用的算法

认证算法/方法	RSA、ECDSA、EdDSA, 预共享密钥(PSK, pre-shared key)
密钥协商算法/方法	(EC)DHE(有限域上或椭圆曲线上的Diffie-Hellman密钥交换协议), PSK, PSK和(EC)DHE混搭方式.
伪随机数发生器	ChaCha20
消息认证码	Poly1305
哈希函数	SHA-256, SHA-384
对称加密算法	AES-128, AES-256

Encryption and decryption: cryptography in the narrow sense

- A pair of mappings: $\mathcal{E} : \mathcal{M} \times \mathcal{K}_e \to \mathcal{C}$ and $\mathcal{D} : \mathcal{C} \times \mathcal{K}_d \to \mathcal{M}$
- Encryption: $\mathcal{E}_{K_e} : \mathcal{M} \to \mathcal{C}$; decryption: $\mathcal{D}_{K_d} : \mathcal{C} \to \mathcal{M}$
- $\mathcal{D}_{K_d}(\mathcal{E}_{K_e}(m)) = m$ for any $m \in \mathcal{M}$
- Encryption is possibly probabilistic (Initialization Vector, Series number, nonce, random number, etc)

Discrete symbols: $(u_1, \cdots, u_n) \in \mathcal{M} = \mathbb{F}_q^n$,
$(k_1, \cdots, k_l) \in \mathcal{K}_e = \mathbb{F}_q^l$, $(z_1, \cdots, z_m) \in \mathcal{C} = \mathbb{F}_q^m$

RSA密码

- 第一个公钥密码（1977）
- **Rivest—Shamir—Adleman**发明
- 教科书方案：给定**n**（**=pq**）和**e**，将明文**m**加密成密文$c = m^e \pmod n$
- 解密：$m = c^d \pmod n$
- 要求：$m = m^{ed} \pmod n$

$$m^{ed-1} = 1 \pmod n$$

- 如何生成（大）素数？
- 如何进行模幂运算（大模数、大幂次）？
- 保证解密正确的原理？
- 解密方能否将模**n**运算化作模**p**和模**q**运算？
- 保证安全性的原理？（大数因子分解的难度？）

一个Windows XP 中的RSA 模数

21335625291600027351142759355194209132914767425698066864818245285802697571587504827160038792867188144217660057955934845800814958268691260056037643469790871613988653520618544234805258949423413033375605873213651488760386443075342912012970548900016706067393246389837569751517347745772076420507479301672647916792373351492517320962556245120580406546060184803670311182370599074873628794261731191112555208060025609009047888480639771734426254325175122847998160609602132860929278043535478577169570898641110787987645625919308715088016517131066837168489289581361754587749922998809128927098697538006934652117684098976045960758751

（十进制617位 = 二进制2048位）

RSA cryptographic system:

- Let p and q be two large prime numbers and $N = pq$
- Let $1 < e, d < N$ such that $ed = 1(\mod \phi(N))$
- Keep (p, q, d) secret (private key) and (N, e) public (public key)
- Encryption: to encrypt a plaintext $m \in \mathbb{Z}_N$, compute

$$c = m^e(\mod N)$$

- Decryption: to decrypt a ciphertext $c \in \mathbb{Z}_N$, compute

$$m = c^d(\mod N)$$

离散对数

- **设G是由g生成的n阶循环群，给定G的元素a，求x，使得**

$$a=g^x,\ 0<x<n$$

- **Z/(11Z)的非零元的集合（n=10），g=2**

- $g^1=2,\ g^2=4,\ g^3=8,\ g^4=16=5,\ g^5=10=-1,$
- $g^6=-2=9,\ g^7=-4=7,\ g^8=14=3,\ g^9=6,\ g^{10}=12=1=g^0$

- $\log_g 1=0,\ \log_g 2=1,\ \log_g 3=8,\ \log_g 4=2,\ \log_g 5=4,$
- $\log_g 6=9,\ \log_g 7=7,\ \log_g 8=3,\ \log_g 9=6,\ \log_g 10=5$

ElGamal密码

- **ElGamal加密方案 (1984)**
  - **B的密钥：私钥x，公钥是 g^x**
  - **加密: A给B发送明文m，密文是**

    $(g^k, m(g^x)^k)$

  （概率加密, 长期、短期密钥）
  - **解密: 设$(u,v)=(g^k,m(g^x)^k)$,**

    B计算:　$v(u^x)^{-1}=m$

- 如何构造一个循环群和它的一个生成元（如何求模p的一个生成元）？
- 如何求模逆？
- 保证安全性的原理？（离散对数的难度？为什么一个256比特的群上的离散对数最多只有128比特的安全强度？为什么最好选素数阶的群？）

DSA数字签名算法

- **DSA是ElGamal签名算法的变种，美国NIST数字签名算法标准**
- **签名：计算 r = (g^k mod p) mod q**

  s = (k^(-1) (H(m) + xr)) mod q

  签名结果是(m, r, s)。
- **签名验证：计算 w = s^(-1)mod q**

u_1 = (H(m) w) mod q， u_2 = (r w) mod q

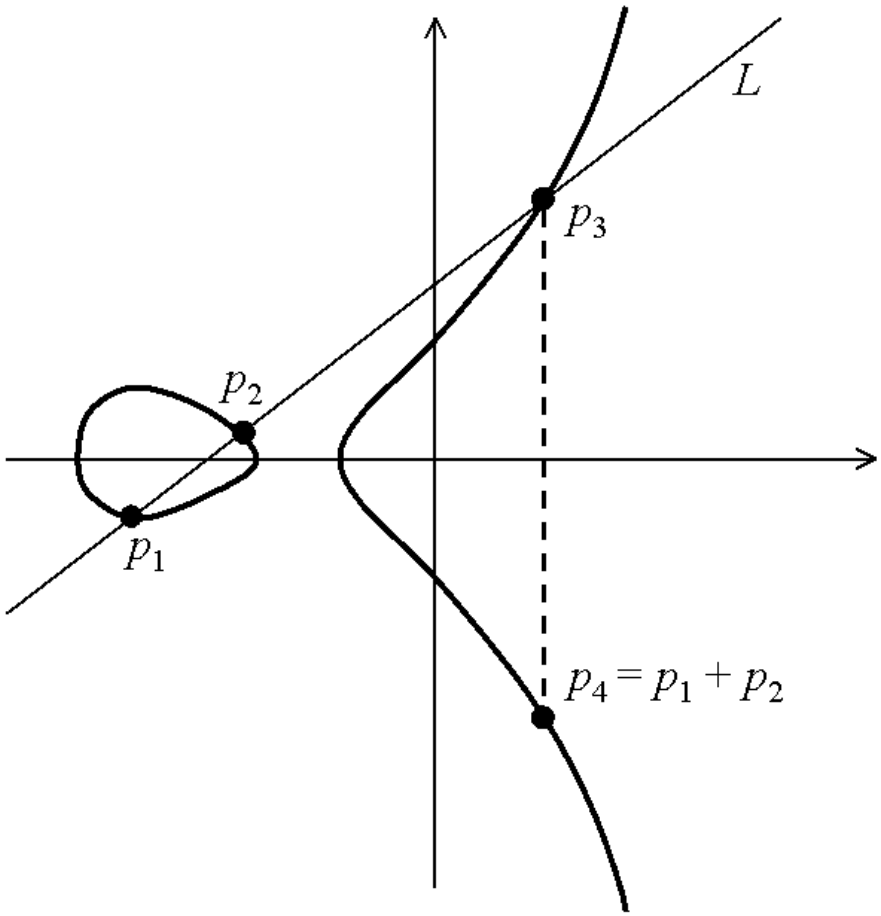
v = ((g^u_1 y^u_2) mod p) mod q

若v = r，则认为签名有效。

- 如何快速计算两个模幂之乘积运算（即 **g^a h^b mod p**）？

椭圆曲线（实数域上）

$y^2=x^3+7x^2+14x+8$



椭圆曲线

- **G=(E/K,+)是加法群**
- **P是基点，阶为素数n（ n 整除G 的阶），私钥为d，公钥为**

  Q=dP=P+…+P (d个)
- **离散对数：用椭圆曲线的点群取代有限域的乘法群**
- **没有发现子指数算法存在**
  - 短密钥

- 点嵌入时的开平方运算？
- 循环群的性质？

Discrete information processing

- No coefficient expansion (pro)
- No measure with mathematical sense
  - Impossible to distinguish conjugate roots (con)
  - No approximation methods, e.g., Newton's iteration (con)
  - Existence of some computational hard problems. e.g., discrete logarithm (pro)

- Symmetric encryption: $K_e = K_d$
- Public key encryption: hard to compute K_d from K_e

- 无限域上数据扩展的例子：实数域上椭圆曲线

因为实数域$\mathbf{R}$的特征不为$2,\ 3$, 所以实数域$\mathbf{R}$上椭圆曲线E的Weierstrass 方程可设为

$$E:\ y^2 = x^3 + a_4x + a_6,$$

其判别式$\Delta = -16(4a_4^3 + 27a_6^2) \neq 0$. E 在$\mathbf{R}$ 上的运算规则为:
设$P_1 = (x_1, y_1),\ P_2 = (x_2, y_2)$是曲线$E$ 上两个点, O 为无穷远点. 则
(1) $O + P_1 = P_1 + O$;
(2) $-P_1 = (x_1, -y_1)$;
(3) 如果$P_3 = (x_3, y_3) = P_1 + P_2 \neq O$,

$$\begin{cases} x_3 = \lambda^2 - x_1 - x_2, \\ \\ y_3 = \lambda(x_1 - x_3) - y_1. \end{cases} \quad \text{其中} \quad \begin{cases} \lambda = \dfrac{y_2 - y_1}{x_2 - x_1} & \text{如果 } x_1 \neq x_2, \\ \\ \lambda = \dfrac{3x_1^2 + a_4}{2y_1} & \text{如果 } x_1 = x_2. \end{cases} \tag{10}$$

例14.2.1 设$P=(0,1)=(x_1,y_1)$是$\mathbf{R}$上椭圆曲线$E:\ y^2=x^3+3x+1$的点. 求$2P=(x_2,y_2),\ 3P=(x_3,y_3),\ 4P=(x_4,y_4),\ 5P=(x_5,y_5),\ 6P=(x_6,y_6),\ 7P=(x_7,y_7),\ 8P=(x_8,y_8)$.

解 根据公式(10), 我们有

$$\lambda_2=\frac{3x_1^2+a_4}{2y_1}=\frac{3}{2},$$

$$x_2=\lambda_2^2-2x_1=\frac{9}{4},\quad y_2=\lambda_2(x_1-x_2)-y_1=\frac{-35}{8}$$

$$\lambda_3=\frac{y_2-y_1}{x_2-x_1}=\frac{-43}{18},$$

$$x_3=\lambda_3^2-x_1-x_2=\frac{280}{81},\quad y_3=\lambda_3(x_1-x_3)-y_1=\frac{5291}{729}$$

$$\lambda_4=\frac{y_3-y_1}{x_3-x_1}=\frac{2281}{1260},$$

$$x_4=\lambda_4^2-x_1-x_3=\frac{-3519}{19600},\quad y_4=\lambda_4(x_1-x_4)-y_1=\frac{-1852129}{2744000}$$

$$\lambda_5 = \frac{y_4 - y_1}{x_4 - x_1} = \frac{510681}{54740},$$

$$x_5 = \lambda_5^2 - x_1 - x_4 = \frac{13333320}{152881}, \quad y_5 = \lambda_5(x_1 - x_5) - y_1 = \frac{-48696013549}{59776471}$$

$$\lambda_6 = \frac{y_5 - y_1}{x_5 - x_1} = \frac{-348255643}{37238058},$$

$$x_6 = \lambda_6^2 - x_1 - x_5 = \frac{2257258249}{9070276644}, \quad y_6 = \lambda_6(x_1 - x_6) - y_1 = \frac{1146658401987805}{863835007021272}$$

$$\lambda_7 = \frac{y_6 - y_1}{x_6 - x_1} = \frac{723333490963}{549812688282},$$

$$x_7 = \lambda_7^2 - x_1 - x_6 = \frac{49390057276560}{333279792955 21},$$

$$y_7 = \lambda_7(x_1 - x_7) - y_1 = \frac{-567521666143702121879}{192403724264235258319}$$

课程内容

- 第一部分 课程概述
- 第二部分 数论算法基础
  - 1）Karatsuba乘法；2）欧几里得算法；3）平方乘算法、Montgomery模幂运算； 4）平方元判定、开平方算法；5）中国剩余定理；6）素数判定算法；7）模素数幂p^n同余式解法
- 第三部分 离散数学基础
  - 代数结构简介，群环域（包括有限域）的基本概念
  - 以数论算法中的具体例子讲解代数结构概念
  - 整数、域上多项式

课程内容

- 第四部分 概率与信息论基础
  - 1）概率的基本概念，条件概率和独立性； 2）随机变量，概率分布，期望值，方差，大数定律，中心极限定理； 3）统计检验； 4）信息的度量，联合熵与条件熵，互信息与平均互信息； 5）纠错编码，汉明码，大数逻辑译码； 6）应用举例：生日悖论及应用，自然语言的冗余性等。
- 第五部分 计算复杂性基础
  - 1）图灵机，算法及其计算复杂度的表示，可计算性；2）复杂性类：归约，P、NP、NPC、BPP等复杂性类，P与NP关系； 3）应用举例：密码学中假设与归约方法

参考用书

- 1） 冯登国等. 信息安全中的数学方法与技术. 清华大学出版社，2009.
- 2） Douglas Stinson. Cryptography-Theory and practice (3ed.). CRC Press, 2006.
- 3）陈恭亮编，《信息安全数学基础》第二版，清华大学出版社，2014年
- 4） 李超等. 信息安全数学基础. 电子工业出版社，2015.
- 5） 裴定一, 徐祥. 信息安全数学基础（第二版）. 人民邮电出版社, 2016.

参考资料:

《SM2椭圆曲线公钥密码算法 第1部分 总则》的资料性附录A、B、C

Henri Cohen (1996). A Course In Computational Algebraic Number Theory. Graduate Texts in Mathematics 138. Springer-Verlag.

Henri Cohen (2000). Advanced Topics in Computational Number Theory. Graduate Texts in Mathematics 193. Springer-Verlag.

课程目的

课程目的：为理解和编码信息安全算法打下基础

核心目标：能够用于计算和快速计算目的的数学理论和算法

工程师加密算法的编程

为什么是这些内容？

- 有限离散集合

- 乘法：Karatsuba快速乘法
- 模逆：欧几里德除法（整数、域上多项式）
  - 辗转相除：求最大公因子（最大公因式）
  - 扩展欧几里德除法——求(s,t)，使得sa+tb=(a,b)。求(s(x),t(x))，使得sf+tg=(f,g)。
    - 求a模素数p的逆。求f(x)模不可约多项式p(x)的逆。
    - 扩展欧几里德除法的复杂度：迭代步数、s和t的规模
    - 秦九韶的大衍求一术
  - BCH纠错码译码中的关键方程求解
- 幂运算g^k
  - 重复平方——相乘法（组合算法，整数模p、多项式模p(x)）
    - 从低位到高位、或从高位到低位
    - 秦九韶算法之多项式赋值
  - Shamir's trick——交换群中两个元素的幂运算$g^k h^l$
    - DSA签名验证
  - 模幂运算的快速方法：Montgomery算法

- 开平方
  - 平方元的判定：基于二次互反律和辗转相除的计算雅可比符号的方法
  - 开平方算法
    - 椭圆曲线的点嵌入
    - 模4余3的素数
- 约减模数：从大模数到小模数
  - 中国剩余定理
    - **RSA—CRT**
  - 模素数幂p^n同余式的解法：幂级数思想
- 素数来自何方：素性判定（**Fermat**、**Solovay-Stassen**、**Miller-Rabin**）

课程目标

- 了解信息安全、电子工程领域中的有限集合及其之间的四则运算、方幂运算和开平方的快速运算原理与算法。

- 了解掌握从无限代数集合到有限代数集合的数论和代数学基本概念和具体实例。了解有限代数集合带来的数据处理优劣性质和密码计算困难问题特性。例如：
  - 离散对数问题的至多平方根复杂度
  - 光滑阶群上的离散对数的脆弱性
  - …

网站

- http://www.shoup.net/ntl/
- **NTL: A Library for doing Number Theory**
- **Victor Shoup's Home Page**
  victor@shoup.net
  Courant Institute，New York University

Q&A