

# 密码学

第六章 哈希函数

网络空间安全学院 朱丹 戚明平 zhudan/mpqi@nwpu.edu.cn

# 知识回顾-椭圆曲线公钥密码

- ✓ 一些国际标准化组织已把椭圆曲线密码作为新的信息安全标准。如,IEEE P1363/D4, ANSI F9.62, ANSI F9.63等标准,分别规范了椭圆曲线密码在 Internet协议安全、电子商务、Web服务器、空间通信、移动通信、智能卡等方面的应用
- ◇ 它密钥短,软件实现规模小、硬件实现节省电路。由于椭圆曲线离散对数问题尚没有发现亚指数算法,所以普遍认为,椭圆曲线密码比RSA和ElGamal密码更安全
- 於 椭圆曲线密码已成为RSA之外呼声最高的公钥密码之一

# 知识回顾-椭圆曲线公钥密码

- グ 160位的椭圆曲线密码的安全性相当于1024位的RSA密码,而且运算速度也较快
- ♪ ElGamal密码建立在有限域GF(p)的乘法群的离散对数问题的困难性之上。而椭圆曲线密码建立在椭圆曲线群的离散对数问题的困难性之上。两者的主要区别是其离散对数问题所依赖的群不同。因此两者有许多相似之处
- ◆ 我国商用密码采用了椭圆曲线密码,并具体颁布了椭圆曲线密码标准算法SM2

- ₹ GF(p)上椭圆曲线密码的基础参数
  - $> T = \langle p, a, b, G, n, h \rangle, p$ 为大于3素数, p确定了有限域GF(p)
  - > 元素 $a, b \in GF(p)$ , a和b确定了椭圆曲线:  $y^2 = x^3 + ax + b$ ,  $a, b \in GF(p)$
  - ightharpoonup G为循环子群 $E_1$ 的生成元,n为素数且为生成元G的阶, G和n确定了循环子群 $E_1$

- ▶ GF(p)上椭圆曲线密码的密钥
  - ▶ 用户的私钥定义为一个随机数d,  $d \in \{1,2,\dots,n-1\}$
  - $\rightarrow$  用户的公钥定义为Q点,Q = dG
  - ightharpoonup 由公钥Q求私钥d是求解椭圆曲线离散对数问题,当p足够大时,这是困难的

- - 學 设d为用户私钥,Q为用户公钥,明文数据为M, 0 ≤ M ≤ n − 1

#### ▶ 加密过程

- $\triangleright$  S1: 选择一个随机数k, 且k ∈ {1, 2, ···, n − 1}
- > S2: 计算点 $X_1(x_1, y_1) = kG$
- > S3: 计算点 $X_2(x_2, y_2) = kQ$ ,如果分量 $x_2 = 0$ ,则转S1
- > S4: 计算密文 $C = Mx_2 \mod n$
- $(X_1, C)$ 为最终的密文数据

# 知识回顾-椭圆曲线公钥密码

- - $\mathcal{L}$  设d为用户私钥,Q为用户公钥,密文数据为 $(X_1, C)$
  - ▶ 解密过程
    - ➤ S1: 利用私钥d求出X<sub>2</sub>(x<sub>2</sub>, y<sub>2</sub>)

$$dX_1 = d(kG) = k(dG) = kQ = X_2$$

> S2: 利用 $X_2(x_2, y_2)$ 计算得到明文M

$$M = Cx_2^{-1} \bmod n$$

# \* 推荐使用256位素域GF(p)上的椭圆曲线 $y^2 = x^3 + ax + b$ , 曲线参数如下:

#### ♪ 密钥:

- 拳 私钥是随机数d, d ∈ [1, n − 1]
- $\checkmark$  公钥 $P = dG = d(x_G, y_G)$

#### ♪ SM2的加密算法:

- F S1: 产生随机数k, 1 ≤ k ≤ n − 1
- ▶ S2: 计算椭圆曲线点 $C_1 = kG = (x_1, y_1)$
- ▶ S3: 计算椭圆曲线点 $kP = (x_2, y_2)$
- F S4: 计算 $t = KDF(x_2||y_2, klen)$ ,若t为全0比特串,则返回S1 //KDF(Z, klen)是密钥派生函数,它利用HASH函数从数据Z产生出长度为klen的密钥数据
- ፆ S5: 计算 $C_2 = M \oplus t$ ;
- F S6: 计算 $C_3 = Hash(x_2||M||y_2)$
- ▶ S7: 输出密文 $C = C_1 || C_2 || C_3 ||$

# 知识回顾-SM2公钥密码算法

#### ♪ SM2的解密算法:

- ▶ S1: 计算 $dC_1 = (x_2, y_2)$
- ▶ S2: 计算椭圆曲线点 $t = KDF(x_2||y_2, klen)$ , 若t为全0比特串,则报错退出
- ▶ S3: 计算椭圆曲线点 $M' = C_2 \oplus t$
- ▶ S5: 计算 $C_3' = Hash(x_2 || M' || y_2)$ , 判断 $C_3' = C_3$ 是否相等(验证)

#### ♪ SM2的算法应用:

- ₹ 我国二代居民身份证采用了SM2椭圆曲线密码
- ₹ 我国的许多商用系统采用了SM2椭圆曲线密码

#### グ 传统ECC与SM2的比较:

#### 传统ECC

- ▶ 计算点 $X_2(x_2, y_2) = kQ$
- ▶ 计算密文 $C = Mx_2 \mod n$
- \* 最终密文数据( $X_1, C$ )

#### SM2

- ▶ 计算点 $t = KDF(x_2||y_2, klen)$
- F 计算密文 $C_2 = M \oplus t$
- ₱ 最终密文数据 $C = C_1 || C_2 || C_3 || C_3 || C_4 || C_5 || C_6 || C_7 || C_8 ||$
- ◈ 传统ECC用 $x_2$ 加密,加密是乘法;SM2处理 $(x_2,y_2)$ 产生t,用t加密,加密是模2加

#### ✔ Hash函数的概念

Arr Hash函数又称哈希函数、杂凑函数、散列函数等,它能够将"任意长度"的消息变为某一固定长度的消息摘要(也称数字指纹、杂凑值、散列值等)。通常记为 h = H(M)或Hash(M)。

#### ✔ Hash函数的性质

- ✔ Hash函数的输入可以是"任意长度"的消息;
- ✔ Hash函数的输出位长固定,多数情况下输入的长度是大于输出的长度的,因此Hash函数具有压缩特性;
- f 有效性,即对给定的f ,计算f =f (f )的运算是高效的;
- 具有极强的错误检测能力,即输入有很小的不同,输出将有很大的不同

#### ✔ Hash函数的定义

₹ Hash函数将任意长的数据M变换为定长的码h, 记为:

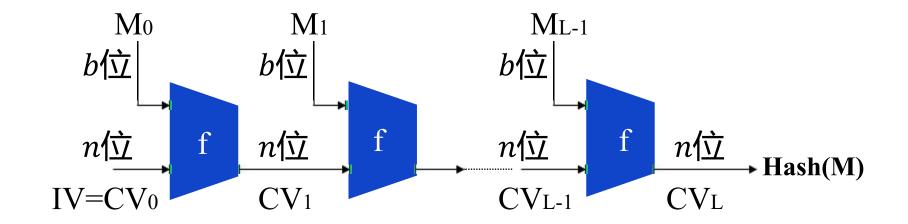
$$h = H(M)$$
或者 $h = Hash(M)$ 

一般,h的长度小于M的长度,因此HASH函数是一种压缩变换

#### ₹ 安全性:

- > 随机性: 哈希函数的输出具有伪随机性
- $\rightarrow$  单向性: 对给定的Hash值h, 找到满足H(x) = h的x在计算上是不可行的
- > 抗弱碰撞性:对任何给定的x,找到满足 $y \neq x$ 且H(x) = H(y)的y在计算上是不可行的
- $\rightarrow$  抗强碰撞性: 找到任何满足H(x) = H(y)的偶对(x,y)在计算上是不可行的

- - $m{\ell}$  Merkle最早提出了Hash函数处理数据M的一般模型,其中 $M=M_0\parallel...\parallel M_{L-1}$



▶ b为分组长度, f为压缩函数, L为链接迭代轮数, n为输出位数。

#### ◆ 安全Hash函数处理数据的一般模型

₹ 分组:将输入M分为L-1个大小为b位的分组

У 填充: 若第L - 1个分组不足b位,则将其填充为b位

▶ 附加:再附加上一个输入的总长度

填充和附加之后, 共L个大小为b位的分组。

由于输入中包含长度,所以攻击者必须找出具有相同Hash值且长度相等的两条报文,或者找出两条长度不等但加入报文长度后Hash值相同的报文,从而增加了攻击的难度。

目前大多数Hash函数均采用这种数据处理模型。

# 章 节安排

Outline



哈希函数简介



SHA-1哈希函数



SM3哈希函数



基于哈希的消息认证码HMAC

# 章节安排

Outline



#### 哈希函数简介



SHA-1哈希函数



SM3哈希函数



基于哈希的消息认证码HMAC

### ♪ 安全哈希函数的应用-认证

- ፟ 报文认证 (方案a)
  - $\triangleright$   $A \rightarrow B$ :  $< M \parallel E(H(M), K) >$
  - $\triangleright$  发送方A生成消息M的Hash码H(M)并使用传统密码对其加密,将加密后的结果附于消息M之后发送给接收方
  - ightharpoonup 由于H(M)受密码保护,所以接收方B可以通过比较H(M)可认证报文的真实性和完整性

# 6.1 哈希函数简介

### ♪ 安全哈希函数的应用-认证

- ₹ 报文认证 (方案b)
  - $\triangleright$   $A \rightarrow B$ :  $< M \parallel H(M \parallel S) >$



# 方案b有什么优点

- $\triangleright$  发送方A生成消息M和秘密值S的Hash码H(M),并将结果附于消息M之后发送给接收方B
- ightharpoonup 假定通信双方AB共享公共的秘密值S, B可以通过验证Hash码来认证数据的真实性和完整性
- ➤ 该方案无需加密,因为秘密值S参与Hash的计算

### ♪ 安全哈希函数的应用-认证和保密

- ₹ 报文认证和保密(方案a)
  - $\triangleright$   $A \rightarrow B$ :  $< E(M \parallel H(M), K) >$
  - $\triangleright$  发送方A生成消息M的Hash码H(M),并将结果附于消息M之后利用密钥加密并将密文发送给接收方B
  - ightharpoonup 由于只有A和B共享秘密密钥,所以B通过比较H(M)可认证报文源和报文的真实性
  - ➤ 由于该方法是对整个报文M和Hash码加密,所以也提供了保密性

### ♪ 安全哈希函数的应用-认证和保密

- 授 报文认证和保密(方案b)
  - $> A \rightarrow B : < E(M \parallel H(M \parallel S), K) >$
  - $\triangleright$  发送方A生成消息M和秘密值S的Hash码H(M),并将结果附于消息M之后利用密钥加密并将密文发送给接收方B
  - ightharpoonup 由于只有A和B共享秘密密钥K以及秘密值S,所以B通过比较 $H(M \parallel S)$  可认证报文源和报文的真实性
  - ➤ 由于该方法是对整个报文M和Hash码加密,所以也提供了保密性

### ◇ 安全哈希函数的应用-认证和数字签名

- 授 报文认证和数字签名
  - $\triangleright$   $A \rightarrow B$ :  $< M \parallel D(H(M), K_{dA}) >$
  - $\triangleright$  发送方A使用公钥密钥用其私钥 $K_{dA}$ 对消息M 的Hash码H (M )签名,并将结果附于消息M 之后发送给接收方B
  - ➤ B可以验证Hash值来认证报文的真实性,因为该方法可提供认证
  - ➤ 由于只有A可以进行签名,所以也提供了数字签名

- ♪ 安全哈希函数的应用-认证、数字签名和保密
  - 授 报文认证、数字签名和保密
    - $\triangleright$   $A \rightarrow B$ :  $< E(M \parallel D(H(M), K_{dA}), K) >$
    - $\triangleright$  发送方A使用公钥密钥用其私钥 $K_{dA}$ 对消息M 的Hash码H (M)签名,并将结果附于消息M 之后用传统密码对M 和签名进行加密,发送给接收方B
    - ➤ B可以验证Hash值来认证报文的真实性,因为该方法可提供认证
    - > 由于只有A可以进行签名,所以也提供了数字签名
    - ➤ 由于该方法是对消息和Hash码的签名整体进行加密,所以也提供了保密性

# 章节安排

Outline



哈希函数简介



SHA-1哈希函数



SM3哈希函数



基于哈希的消息认证码HMAC

- ◆ 安全哈希函数的设计同分组密码一样复杂困难,一些国际组织和密码学者不断制定和颁布Hash函数标准,包括
  - Rivest—MD4, MD5, MD6
  - NIST——SHA-0, SHA-1, SHA-2, SHA-3

  - ▶ 澳大利亚——HAVAL
  - **▶** 中国——SM3

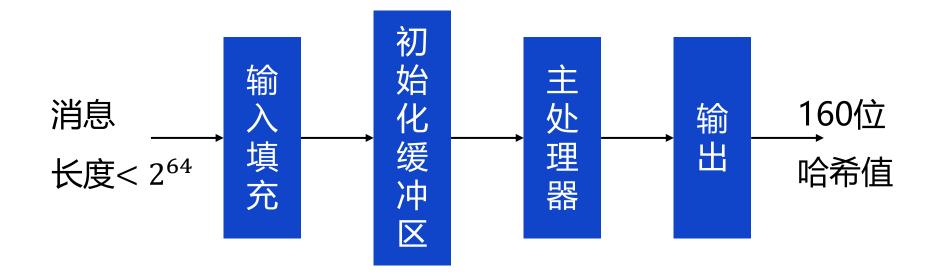
- ◆ SHA系列哈希函数:
  - ▼ SHA系列Hash函数由美国标准与技术研究所(NIST)设计
  - ▶ 1993年公布了SHA-0 (FIPS PUB 180), 后来发现它不安全;
  - ▶ 1995年又公布了SHA-1 (FIPS PUB 180-1); 【2017年Google给出第一个碰撞】
  - № 2002年又公布了SHA-2 (FIPS PUB 180-2), SHA-2包括3个Hash算法: SHA-256, SHA-384, SHA-512; 【2008年补充了SHA-224】
  - № 2005年, 王小云院士给出了一种攻击SHA-1的方法, 用2<sup>69</sup>次操作找到一个强碰撞, 以前认为是穷举(生日) 攻击2<sup>80</sup>次操作
  - ▶ NIST于2007年公开征集SHA-3,并于2012公布SHA-3获胜算法为Keccak

# 6.2 SHA-1哈希函数

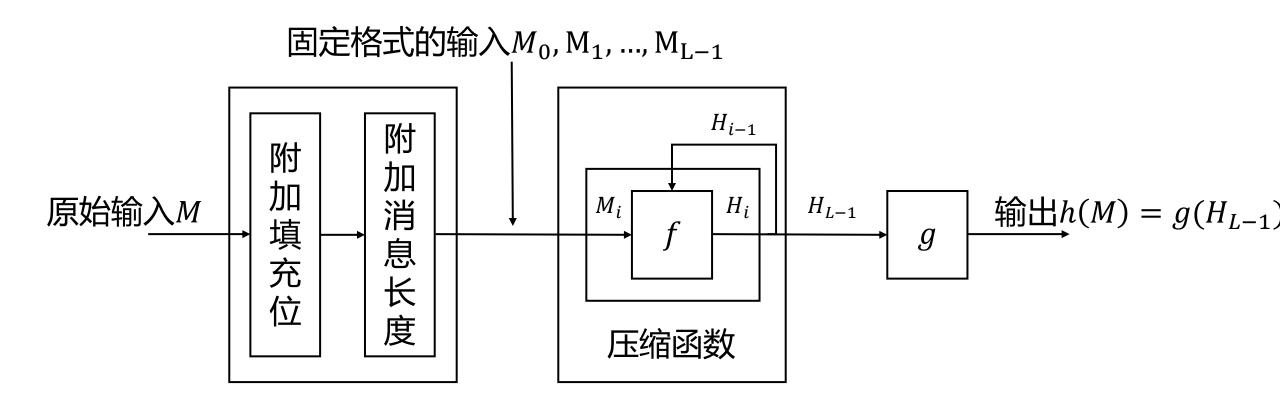
# ♪ 各版本SHA算法参数比较:

	SHA-1	SHA-224	SHA-256	SHA-384	SHA-512
消息摘要长度	160	224	256	384	512
消息长度	< 2 <sup>64</sup>	< 2 <sup>64</sup>	< 2 <sup>64</sup>	< 2 <sup>128</sup>	< 2 <sup>128</sup>
分组长度	512	512	512	1024	1024
字长度	32	32	32	64	64
步骤数	80	64	64	80	80
穷尽安全性	80	112	128	192	256

❖ SHA-1的结构: 采用Merkle提出的安全Hash模型



♪ SHA-1算法

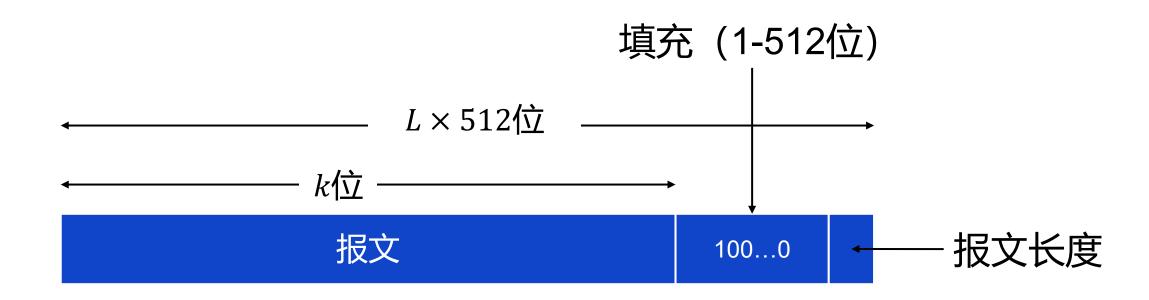


- ♪ SHA-1算法的输入填充
  - ✔ 目的是使填充后的报文长度满足:

长度=448 mod 512

- ▶ 填充方法是在报文后附加一个1和若干个0
- > 然后附上表示填充前报文长度的64位数据(最高有效位在前)
- 孝 若报文本身已经满足上述要求,仍然需要填充(例如,若报文长度为448位,则仍需填充512位使其长度为960位),因此填充位数在1到512位之间
- 學 经过填充和附加后,数据的长度为512位的整数倍

♪ SHA-1算法的输入填充



对消息进行填充, 使之加上64位长度信息后, 总长度为512的整数倍

- ♪ SHA-1算法的初始化缓冲区
  - **ᢞ** 缓冲区由5个32位的寄存器 (A, B, C, D, E) 组成,用于保存160位的中间结果和最终结果
  - 學 将寄存器初始化为下列32位的整数:

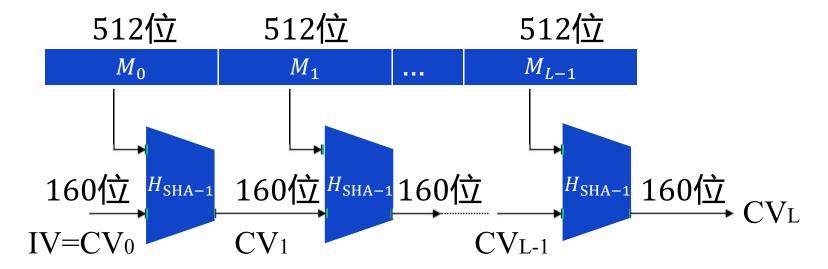
A: 67452301 B: EFCDAB89

C: 98BADCFE D: 10325476

E: C3D2E1F0

注意: 高有效位存于低地址

- ♪ SHA-1算法的主处理
  - ₹ 主处理是SHA-1 HASH函数的核心
  - ∮ 每次处理一个512位的分组,链接迭代处理(填充后报文)的所有L个分组数



利用SHA-1算法产生报文摘要 $CV_L$ ,其中 $H_{SHA-1}$ 是SHA-1算法的压缩函数

- グ SHA-1算法的主处理
  - 学 将填充后的消息M按512比特进行分组,记为  $M = M_0 \parallel M_1 \parallel ... \parallel M_{L-1}$ ,其中L = (K + n + 65)/512,n为填充时插入的0的个数
  - ▶ 对分组进行迭代压缩处理

$$\begin{cases} CV_0 = IV, IV$$
为预设的初始值 (160 bit)   
  $CV_{i+1} = H_{SHA-1}(CV_i, M_i), 0 \le i \le L-1$    
  $H(M) = CV_L$ 

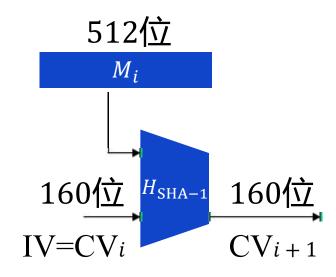
A = 0x67452301

B=0xEFCDAB89

C=0x98BADCFE

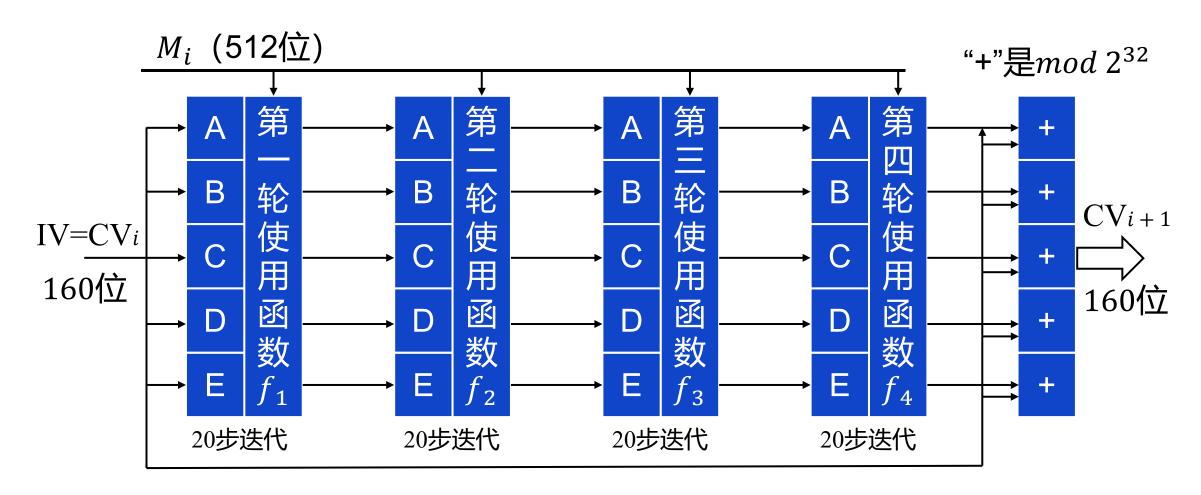
D=0x10325476

E=0xC3D2E1F0



- グ SHA-1算法的主处理
  - ▶ 压缩函数是主处理的核心
  - ▶ 它由四层运算(每层迭代20步)组成,四层的运算结构相同
  - 争较的输入是当前要处理的512位的分组 $M_i$ 和160位缓冲区ABCDE的内容,每轮都对ABCDE的内容更新,而且<mark>每轮使用的逻辑函数不同</mark>,分别为 $f_1$ , $f_2$ , $f_3$ 和 $f_4$
  - 第四轮的输出与第一轮的输入相加得到压缩函数的输出

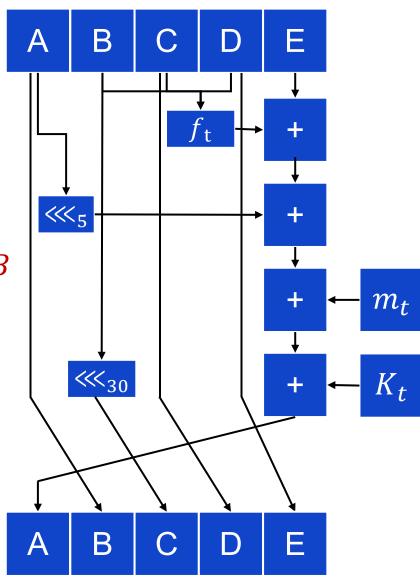
#### ♪ SHA-1算法的主处理



- グ SHA-1算法的主处理
  - ₹ SHA-1算法压缩函数中的单步操作
    - $\triangleright (A, B, C, D, E)$

$$= ((E + f_t(B, C, D) + (A \ll 5) + m_t + K_t, A, B \ll 30, C, D)$$

逻辑函数:  $f_1 = (B \land C) \lor (\neg B \land D)$ ,  $f_2 = B \oplus C \oplus D$ ,  $f_3 = (B \land C) \lor (B \land D) \lor (C \land D)$ ,  $f_4 = B \oplus C \oplus D$ 



## グ SHA-1算法的主处理

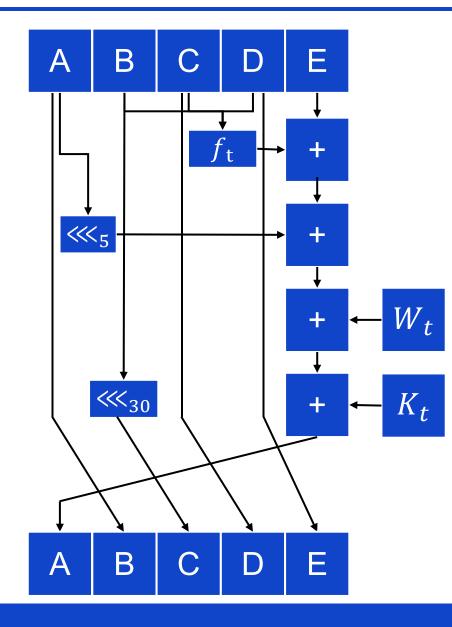
- ₹ SHA-1算法压缩函数中的单步操作
  - 》常量:  $K_t = 0x5A827999$   $0 \le t \le 19$   $K_t = 0x6ED9EBA1 \quad 20 \le t \le 39$

$$K_t = 0x8F1BBCDC$$
  $40 \le t \le 59$ 

$$K_t = 0xCA62C1D6$$
  $60 \le t \le 79$ 

> 512位数据分组扩展:

当
$$0 \le t \le 15$$
时, $W_t = m_t$ ;当 $16 \le t \le 79$ 时,
$$m_t = (m_{t-16} \oplus m_{t-14} \oplus m_{t-8} \oplus m_{t-3}) \ll 1$$



# 章节安排

Outline



哈希函数简介



SHA-1哈希函数



SM3哈希函数



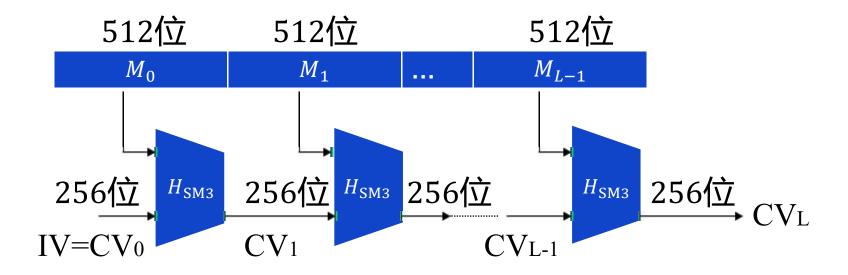
基于哈希的消息认证码HMAC

- - ₹ 2010年12月国家密码管理局正式颁布
  - ▶ 适用于商用密码应用中的数字签名和验证、消息验证码的生成与验证以及随机数的生成
  - ▼ 可满足多种密码应用的安全需求
  - ▶ 面向32bit的字设计,数据分组长度为512 bit,输出Hash值位长为256 bit
  - ₹ SM3标准文档

http://www.oscca.gov.cn/sca/xxgk/2010-12/17/1002389/files/302a3ada057c4a73830536d03e683110.pdf

- グ SM3哈希函数的"压缩函数"+"迭代结构"
  - ▶ 与SHA-1哈希函数相同

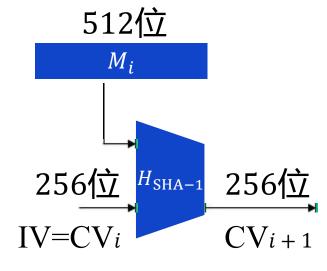




- ♪ SM3哈希函数的 "压缩函数" + "迭代结构"
- グ 与SHA-1哈希函数相同
  - 》将填充后的消息M按512比特进行分组,记为  $M = M_0 \parallel M_1 \parallel ... \parallel M_{L-1}$ ,其中L = (K + n + 65)/512,n为填充时插入的0的个数
  - ▶ 对分组进行迭代压缩处理

$$\begin{cases} CV_0 = IV, IV$$
为预设的初始值 (256 bit)、  $CV_{i+1} = H_{SHA-1}(CV_i, M_i), 0 \le i \le L-1 \ H(M) = CV_L \end{cases}$ 

A=0x7380166f B=0x4914b2b9 C=0x172442d7 D=0xda8a0600 E=0xa96f30bc F=0x163138aa G=0xe38dee4d H=0xb0fb0e4e



### SM3压缩函数: $ABCDEFGH \leftarrow CV^{(i)}$ FOR i = 0 TO 63 $SS1 \leftarrow ((A <<< 12) + E + (T_j <<< j)) <<< 7;$ $SS2 \leftarrow SS1 \oplus (A <<< 12);$ $TT1 \leftarrow FF_i(A, B, C) + D + SS2 + W'_i$ $TT2 \leftarrow GG_i(E, F, G) + H + SS1 + W_i$ $D \leftarrow C$ $C \leftarrow B <<< 9$ $B \leftarrow A$ $A \leftarrow TT_1$ $H \leftarrow G$ $G \leftarrow F <<< 19$ : $F \leftarrow E$ $E \leftarrow P_0(TT2)$ **ENDFOR** $V^{(i+1)} \leftarrow ABCDEFGH \oplus V^{(i)}$

## 逻辑函数: (提供混淆作用) $FF_{j}(X,Y,Z) = \begin{cases} X \oplus Y \oplus Z & 0 \le j \le 15 \\ (X \wedge Y) \vee (X \wedge Z) \vee (Y \wedge Z) & 16 \le j \le 63 \end{cases}$ $GG_{j}(X,Y,Z) = \begin{cases} X \oplus Y \oplus Z & 0 \le j \le 15 \\ (X \wedge Y) \vee (\neg X \wedge Z) & 16 \le j \le 63 \end{cases}$

置换函数: (提供扩散作用)
$$P_0(X) = X \oplus (X <<< 9) \oplus (X <<< 17)$$

$$P_1(X) = X \oplus (X <<< 15) \oplus (X <<< 23)$$

式中X,Y,Z为32位字,符号a <<< n表示把a循环左移n位。

### 常量:

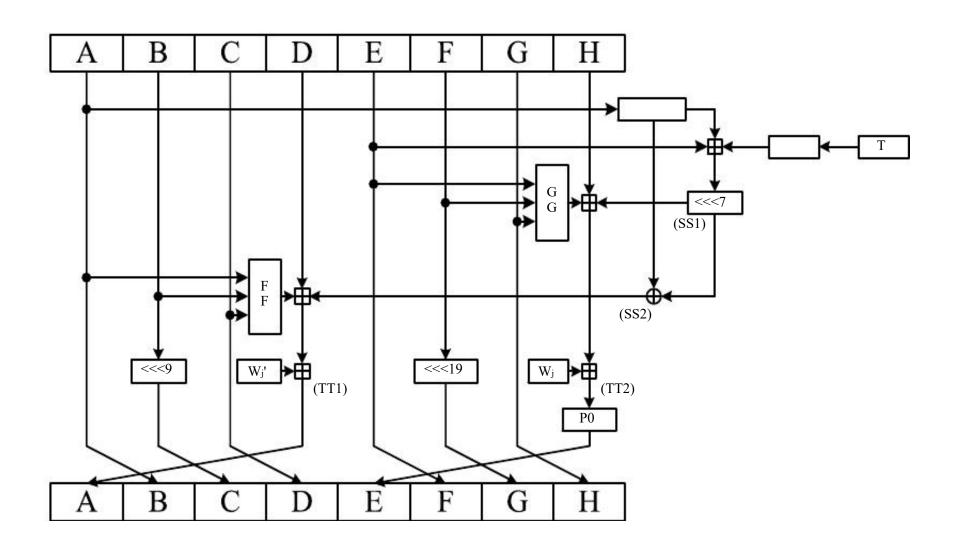
$$T_{j=} \begin{cases} 79cc45190 & 0 \le j \le 15 \\ 7a879d8a & 16 \le j \le 63 \end{cases}$$

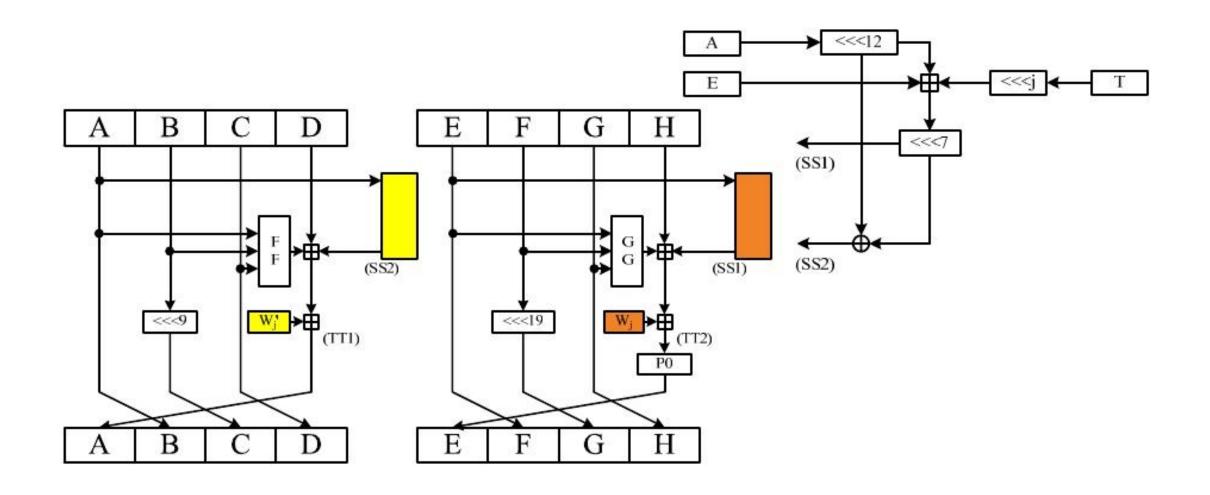
### 512位数据分组扩展:

$$W_{j} = m_{j}$$
  $0 \le j \le 15$ ,即512 bit分组划分成16个字  $W_{j} \leftarrow P_{1}(W_{j-16} \oplus W_{j-9} \oplus (W_{j-3} <<<15)) \oplus W_{j-13} <<<<7) \oplus W_{j-6}$ 

 $16 \le j \le 67$ 

$$W_j' = W_j \oplus W_{j+4} \qquad 0 \le j \le 63$$





- グ SM3哈希函数
  - 我国商用密码杂凑算法标准
  - ፟ 使用M-D结构
  - **※** 采用双路消息扩展输入
  - ₱ 非对称Feistel结构
  - **௺** 安全性只有经过实践检验,才能给出正确结论

# 章节安排

Outline



哈希函数简介



SHA-1哈希函数



SM3哈希函数



基于哈希的消息认证码HMAC

## 6.4 基于哈希的消息认证码HMAC

- ✔ HMAC是密钥相关的哈希运算消息认证码,是 一种基于Hash函数和密钥进行消息认证的方法
- ♪ HMAC的构造:
  - 基于分组密码算法构造
  - ፟ 基于Hash算法构造 (HMAC)
- ◆ HMAC的作用:
  - 斧 消息完整性认证
  - **斧** 信源身份认证

