信息安全导论

第三章 密码学基础

黄 玮

温故

- 信息系统安全要素
- 网络安全基础
- 信息安全保障体系
- 信息安全技术框架

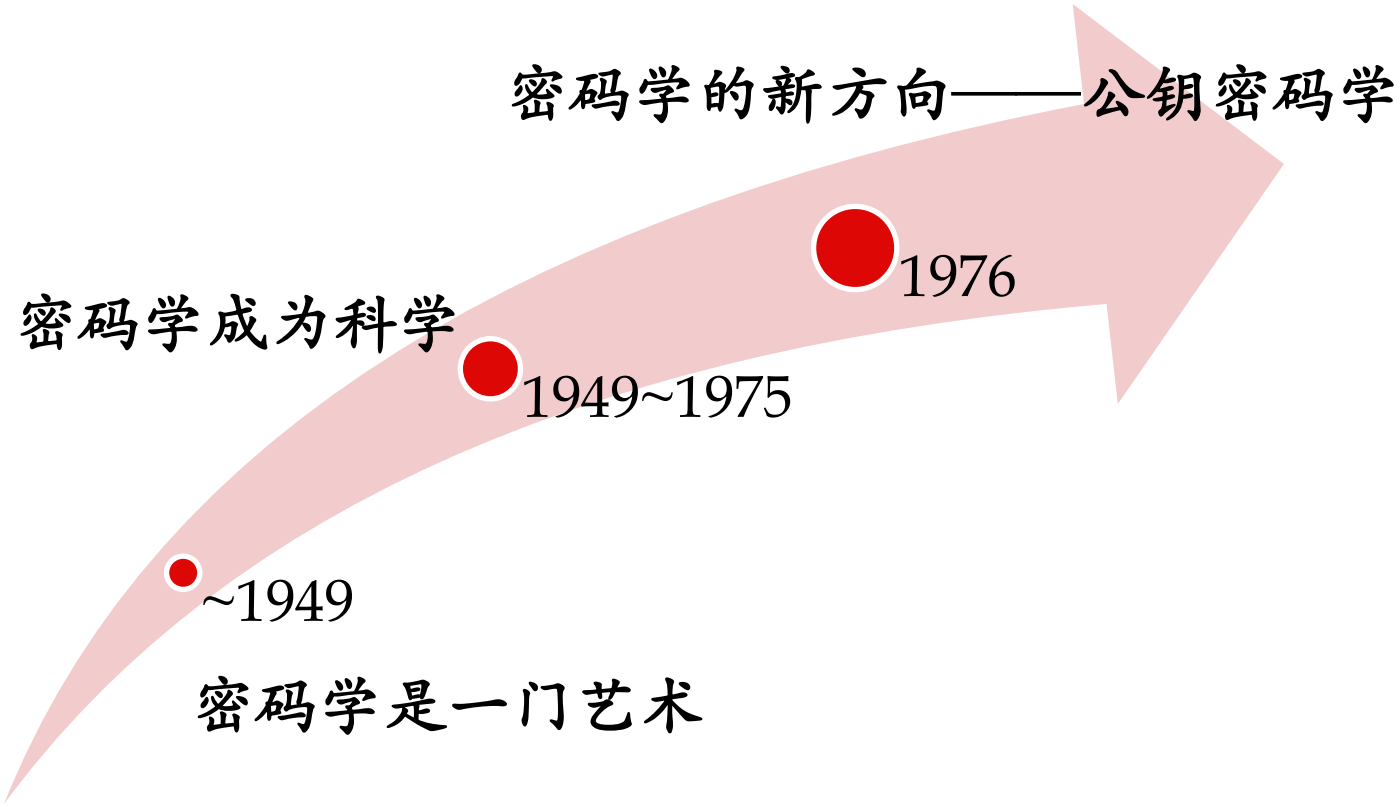
知新

- 密码学简史
- 密码学基本概念
- 流密码
- 分组密码

本章内容提要

- 密码学简史
- 密码学基本概念
- 流密码
- 分组密码

密码学发展阶段



第1阶段—古典密码

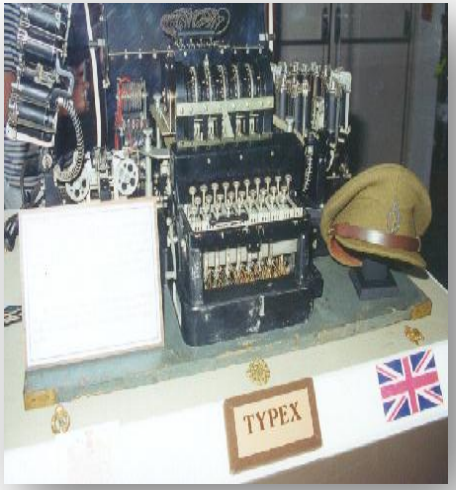
- 密码学还不是科学,而是艺术
- 出现一些密码算法和加密设备
- 密码算法的基本手段出现，针对的是字符
- 简单的密码分析手段出现
- 主要特点：数据的安全基于算法的保密

第1阶段—古典密码

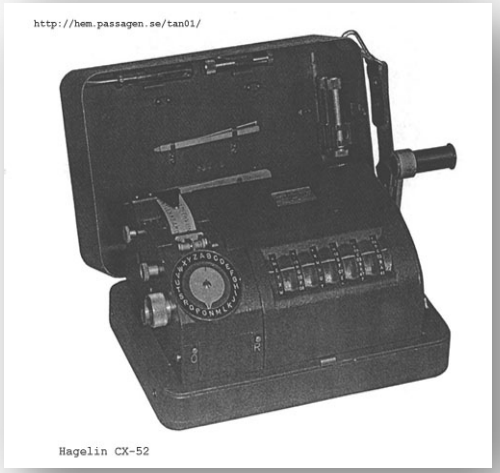
Phaistos圆盘，一种直径约为160mm的Cretan-Mnoan粘土圆盘，始于公元前17世纪。表面有明显字间空格的字母，至今还没有破解。

20世纪早期密码机


ENIGMA
© Deutsches Museum


TYPEX


TUNNY


http://hem.passagen.se/tan01/
Hagelin CX-52


KRYHA-CHIFFRIERMASCHINEN
MODELL: COMMERCIAL

中途岛战役

- 约瑟夫·罗谢福特: 一位优秀的天才般的密码专家。1940年，他帮助破解了日本海军的行动代码JN-25
- 正是由于得到了他的关于日军企图攻占莫尔兹比港口的报告，美国海军才派遣第十七特遣舰队参加了珊瑚海的战斗

中途岛战役

- 1942年5月中旬，罗谢福特发现一支敌军主力部队将要展开行动
- 日军无线电波发射频繁；这些电波表明日军正在计划大规模的行动。但是攻击的目标是哪儿？
- 通过研究破译的密码，罗谢福特注意到，日军反复使用了“AF”这两个字母。无论AF代表什么意思，罗谢福特猜测，它就是日军攻击的目标。

中途岛战役

- 罗谢福特最后认为，AF一定是中途岛

- 检验: 得到尼米兹海军上将的准许之后，向中途岛发送一个伪造的情报--很清楚地报告了岛上蒸馏厂的倒闭。两天后，他们截获到一个新的日军报告，说AF缺少淡水

纳瓦霍语

- 1942至1945年太平洋战争期间，美国海军陆战队征召了420名纳瓦霍族人，让他们用自己的土著语言编制和传递密码
- 由于纳瓦霍语没有文字，语法和发音又极其复杂，当时与美军作战的日军一直无法破译，因此这种密码也被称为“不可破译的密码”

第1阶段—古典密码

- 1883年Kerchoffs第一次明确提出了编码的原则

  —加密算法应建立在算法的公开不影响明文和密钥的安全。

  —这一原则已得到普遍承认，成为判定密码强度的衡量标准，实际上也成为传统密码和现代密码的分界线。

第2阶段—1949~1975

- 计算机使得基于复杂计算的密码成为可能
- 相关技术的发展
  - —1949年Shannon的“The Communication Theory of Secret Systems”
  - —1967年David Kahn的《The Codebreakers》
  - —1971-73年IBM Watson实验室的Horst Feistel等几篇技术报告
- 主要特点：数据的安全基于密钥而不是算法的保密

第3阶段—1976~

- 1976年：Diffie & Hellman 的“New Directions in Cryptography”提出了不对称密钥
- 1977年Rivest, Shamir & Adleman提出了RSA公钥算法
- 90年代逐步出现椭圆曲线等其他公钥算法
- 主要特点：公钥密码使得发送端和接收端无密钥传输的保密通信成为可能

第3阶段—1976~

- 1977年DES正式成为标准
- 80年代出现“过渡性”的“Post DES”算法,如IDEA,RCx,CAST等
- 90年代对称密钥密码进一步成熟 Rijndael,RC6, MARS, Twofish, Serpent等出现
- 2001年Rijndael成为DES的替代者

本章内容提要

- 密码学简史
- 密码学基本概念
- 流密码
- 分组密码

基本概念

- 密码学(Cryptology): 是研究信息系统安全保密的科学.
- 密码编码学(Cryptography): 主要研究对信息进行编码,实现对信息的隐蔽.
- 密码分析学(Cryptanalysis):主要研究加密消息的破译或消息的伪造.

基本概念

明文（Plaintext）：消息的初始形式；

密文（Ciphertext）：加密后的形式

记：

明文记为P且P为字符序列，$P=[P1,P2,\cdots,Pn]$

密文记为C，$C=[C1,C2,\cdots,Cn]$

明文和密文之间的变换记为 $C=E(P)$及$P=D(C)$

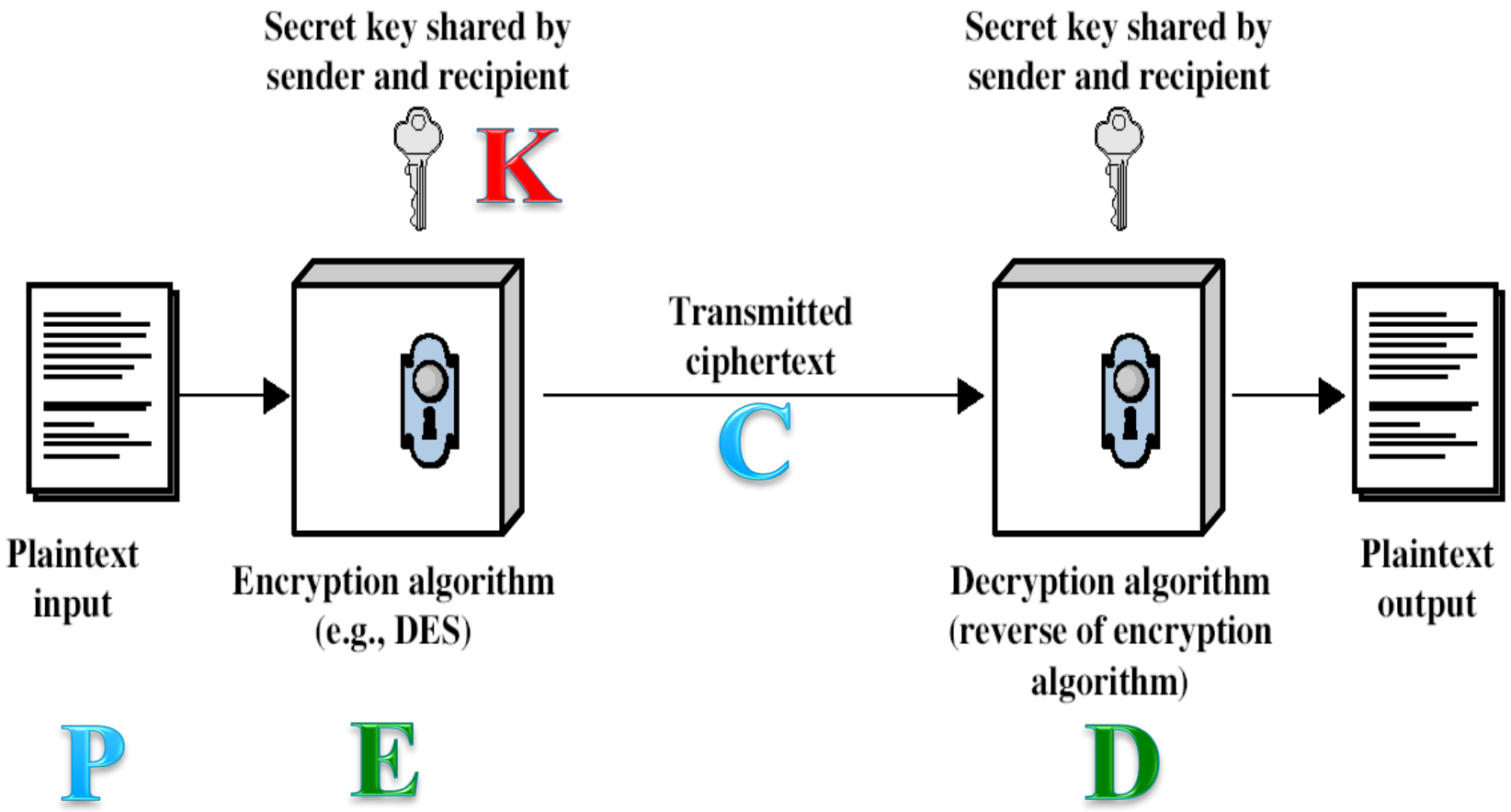
其中 C表示密文，E为加密算法；P为明文，D为解密算法

我们要求密码系统满足：$P=D(E(P))$

基本概念

- 需要密钥的加密算法，记为：C=E(K,P),即密文消息同时依赖于初始明文和密钥的值。实际上，E是一组加密算法，而密钥则用于选择其中特定的一个算法。
- 加密与解密的密钥相同，即：P=D(K,E(K,P))
- 加密与解密的密钥不同，则：$P=D(K_D,E(K_E,P))$

常规加密的简化模型——五元组模型



常规加密的安全性分析

- 加密算法足够强大
  - —仅知密文很难破译出明文
- 基于密钥的安全性，而不是基于算法的安全性
  - —基于密文和加/解密算法很难破译出明文
- 算法开放性
  - —开放算法，便于实现

密码编码系统分类

- 保密内容
- 密钥数量
- 明文处理的方式

保密内容

- 受限制的（restricted)算法

  —算法的保密性基于保持算法的秘密

- 基于密钥（key based)的算法

  —算法的保密性基于对密钥的保密

密钥数量(1/2)

- 对称密钥算法（symmetric cipher）

  —加密密钥和解密密钥相同，或实质上等同，即从一个易于推出另一个

  —又称秘密密钥算法或单密钥算法

密钥数量(2/2)

- 非对称密钥算法（asymmetric cipher）
  - —加密密钥和解密密钥不相同，从一个很难推出另一个
  - —又称公开密钥算法（public-key cipher）
  - —公开密钥算法用一个密钥进行加密，而用另一个进行解密
  - —其中的加密密钥可以公开，又称公开密钥（public key），简称公钥。解密密钥必须保密，又称私人密钥（private key）私钥，简称私钥

明文处理方式

- 分组密码（block cipher）

  —将明文分成固定长度的组，用同一密钥和算法对每一块加密，输出也是固定长度的密文。

- 流密码（stream cipher）

  —又称序列密码。序列密码每次加密一位或一字节的明文。

密码分析

- 试图破译单条消息

  —密码攻击

- 试图识别加密的消息格式，以便借助直接的解密算法破译后续的消息
- 试图找到加密算法中的普遍缺陷（无须截取任何消息）

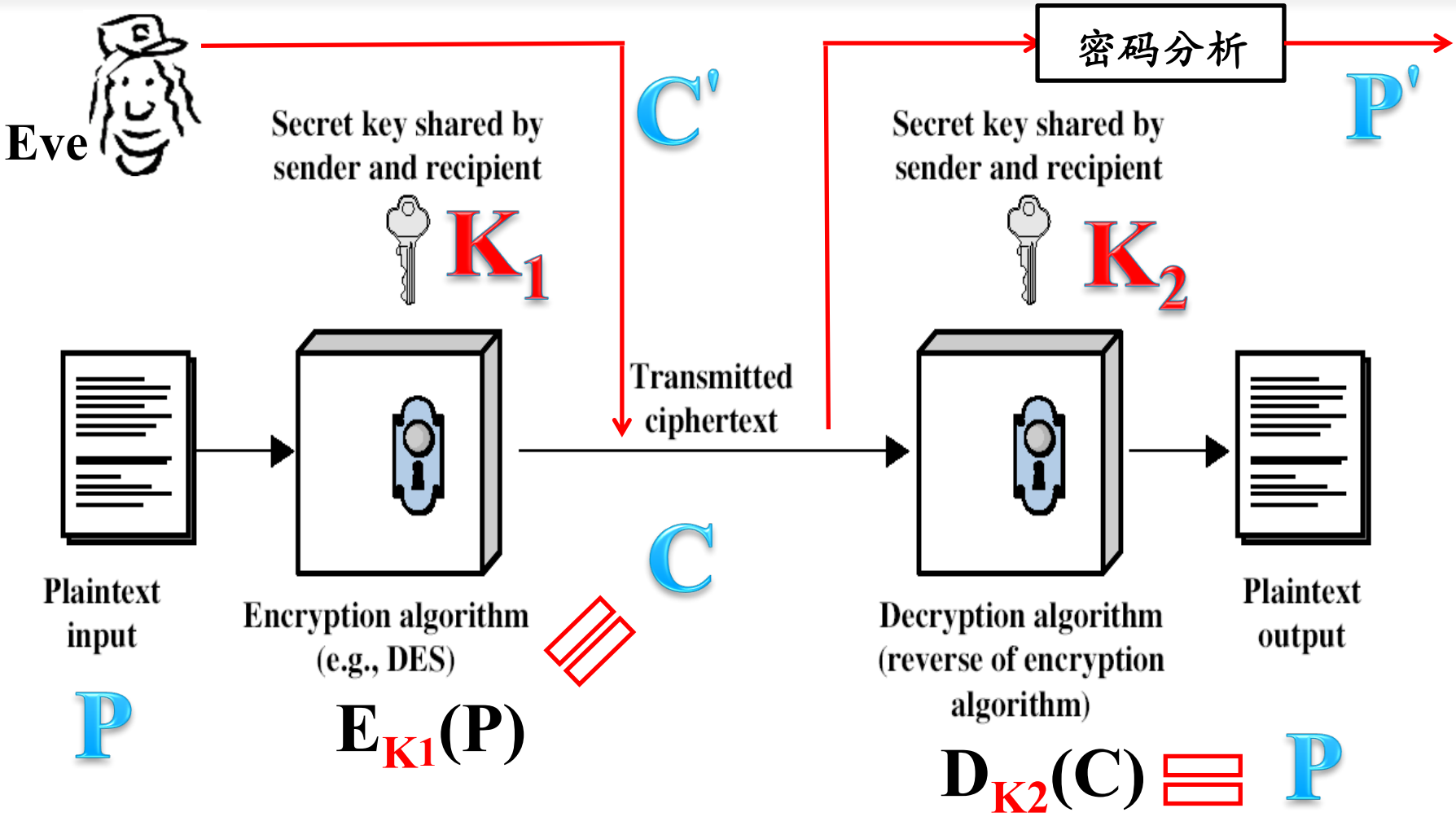
密码分析的条件与工具

- 穷举攻击
  - —已知加密算法
  - —计算机实现漏洞
- 统计分析
  - —截取到明文、密文中已知或推测的数据项
  - —语言特性
- 数学分析
  - —加密算法设计缺陷
- 技巧与运气

密码分析类型

攻击类型	密码破译者已知的东西
唯密文	● 加密算法 ● 待破译的密文
已知明文	● 加密算法 ● 待破译的密文 ● 由密钥形成的一个或多个明文－密文对
选择明文	● 加密算法 ● 待破译的密文 ● 由破译者选择的明文消息，连同对应的由密钥生成的密文
选择密文	● 加密算法 ● 待破译的密文 ● 由破译者选择的猜测性密文，连同它对应的由密钥生成的已破译明文
选择文本	● 加密算法 ● 待破译的密文 ● 由破译者选择的明文消息，连同对应的由密钥生成的密文 ● 由破译者选择的猜测性密文，连同它对应的由密钥生成的已破译明文

保密通信系统模型——威胁建模



加密方案的安全性

- 无条件安全：无论提供的密文有多少，如果由一个加密方案产生的密文中包含的信息不足以唯一地决定对应的明文
- 除了一次一密的方案外，没有无条件安全的算法
- 安全性体现在：

  —破译的成本超过加密信息的价值

  —破译的时间超过该信息有用的生命周期

攻击的复杂性分析

- 数据复杂性（data complexity）用作攻击输入所需要的数据
- 处理复杂性（processing complexity）完成攻击所需要的时间
- 存储需求（storage requirement）进行攻击所需要的数据量

密钥搜索所需平均时间

密钥长度 (bit)	密钥数量	每微秒加密1次所需时间	每微秒加密100万次所需时间
32	$2^{32}=4.3\times10^{9}$	2^{31}us=35.8 minutes	2.15ms
56	$2^{56}=7.2\times10^{16}$	2^{55}us=1142years	10.01 hours
128	$2^{128}=3.4\times10^{38}$	2^{127}us=5.4×10^{24} years	5.4×10^{18}years
26 characters (permutation)	$26!=4\times10^{26}$	2×10^{26}us=6.4×10^{12} years	6.4×10^{6}years

文本加密与二进制加密

- 计算机处理的所有数据都是二进制
  —只有0和1
- 比特（bit / b）
  —计算机表示数据的最小单位
- 字节（byte / B）
  —1字节=8比特
- 现代密码与古典密码都是以二进制为基础

字符编码与计算机数据存储

Ctrl	十进制	十六进制	字符	代码
^@	0	00		NUL
^A	1	01		SOH
^B	2	02		STX
^C	3	03		ETX
^D	4	04		EOT
^E	5	05		ENQ
^F	6	06		ACK
^G	7	07		BEL
^H	8	08		BS
^I	9	09		HT
^J	10	0A		LF
^K	11	0B		VT
^L	12	0C		FF
^M	13	0D		CR
^N	14	0E		SO
^O	15	0F		SI
^P	16	10		DLE
^Q	17	11		DC1
^R	18	12		DC2
^S	19	13		DC3
^T	20	14		DC4
^U	21	15		NAK
^V	22	16		SYN
^W	23	17		ETB
^X	24	18		CAN
^Y	25	19		EM
^Z	26	1A		SUB
^[	27	1B		ESC
^\	28	1C		FS
^]	29	1D		GS
^^	30	1E	▲	RS
^-	31	1F	▼	US

十进制	十六进制	字符
32	20	
33	21	!
34	22	"
35	23	#
36	24	$
37	25	%
38	26	&
39	27	'
40	28	(
41	29	)
42	2A	*
43	2B	+
44	2C	,
45	2D	-
46	2E	.
47	2F	/
48	30	0
49	31	1
50	32	2
51	33	3
52	34	4
53	35	5
54	36	6
55	37	7
56	38	8
57	39	9
58	3A	:
59	3B	;
60	3C	<
61	3D	=
62	3E	>
63	3F	?

十进制	十六进制	字符
64	40	@
65	41	A
66	42	B
67	43	C
68	44	D
69	45	E
70	46	F
71	47	G
72	48	H
73	49	I
74	4A	J
75	4B	K
76	4C	L
77	4D	M
78	4E	N
79	4F	O
80	50	P
81	51	Q
82	52	R
83	53	S
84	54	T
85	55	U
86	56	V
87	57	W
88	58	X
89	59	Y
90	5A	Z
91	5B	[
92	5C	\
93	5D	]
94	5E	^
95	5F	_

十进制	十六进制	字符
96	60	`
97	61	a
98	62	b
99	63	c
100	64	d
101	65	e
102	66	f
103	67	g
104	68	h
105	69	i
106	6A	j
107	6B	k
108	6C	l
109	6D	m
110	6E	n
111	6F	o
112	70	p
113	71	q
114	72	r
115	73	s
116	74	t
117	75	u
118	76	v
119	77	w
120	78	x
121	79	y
122	7A	z
123	7B	{
124	7C	|
125	7D	}
126	7E	~
127	7F	⌂

十进制	十六进制	字符
128	80	Ç
129	81	ü
130	82	é
131	83	â
132	84	ä
133	85	à
134	86	å
135	87	ç
136	88	ê
137	89	ë
138	8A	è
139	8B	ï
140	8C	î
141	8D	ì
142	8E	Ä
143	8F	Å
144	90	É
145	91	æ
146	92	Æ
147	93	ô
148	94	ö
149	95	ò
150	96	û
151	97	ù
152	98	ÿ
153	99	Ö
154	9A	Ü
155	9B	¢
156	9C	£
157	9D	¥
158	9E	₧
159	9F	ƒ

十进制	十六进制	字符
160	A0	á
161	A1	í
162	A2	ó
163	A3	ú
164	A4	ñ
165	A5	Ñ
166	A6	ª
167	A7	º
168	A8	¿
169	A9	⌐
170	AA	¬
171	AB	½
172	AC	¼
173	AD	¡
174	AE	«
175	AF	»
176	B0	░
177	B1	▒
178	B2	▓
179	B3	│
180	B4	┤
181	B5	╡
182	B6	╢
183	B7	╖
184	B8	╕
185	B9	╣
186	BA	║
187	BB	╗
188	BC	╝
189	BD	╜
190	BE	╛
191	BF	┐

十进制	十六进制	字符
192	C0	└
193	C1	┴
194	C2	┬
195	C3	├
196	C4	─
197	C5	┼
198	C6	╞
199	C7	╟
200	C8	╚
201	C9	╔
202	CA	╩
203	CB	╦
204	CC	╠
205	CD	═
206	CE	╬
207	CF	╧
208	D0	╨
209	D1	╤
210	D2	╥
211	D3	╙
212	D4	╘
213	D5	╒
214	D6	╓
215	D7	╫
216	D8	╪
217	D9	┘
218	DA	┌
219	DB	▉
220	DC	▄
221	DD	▌
222	DE	▐
223	DF	▀

十进制	十六进制	字符
224	E0	α
225	E1	ß
226	E2	Γ
227	E3	π
228	E4	Σ
229	E5	σ
230	E6	µ
231	E7	τ
232	E8	Φ
233	E9	Θ
234	EA	Ω
235	EB	δ
236	EC	∞
237	ED	φ
238	EE	ε
239	EF	∩
240	F0	≡
241	F1	±
242	F2	≥
243	F3	≤
244	F4	⌠
245	F5	⌡
246	F6	÷
247	F7	≈
248	F8	°
249	F9	∙
250	FA	·
251	FB	√
252	FC	n
253	FD	2
254	FE	■
255	FF	

* ASCII 代码 127 拥有代码 DEL。在 MS-DOS 下，此代码具有与 ASCII 8 (BS) 相同的效果。
DEL 代码可由 CTRL + BKSP 键生成。

计算机数据处理

- 逻辑和（或/并集）
  —A+B
- 逻辑积（与/交集）
  —A*B
- 逻辑否
  —!A
- 否定逻辑和
  —!(A+B)
- 否定逻辑积
  —!(A*B)
- 逻辑异或（XOR运算）
  —!A*B+A*!B=A XOR B

异或运算在密码学中的应用

- 假设（1101）是明文，（1001）是加密密钥

—(1101) XOR (1001)　　=　　(0100)

—明文　　加密钥匙　　密文

—(0100) XOR (1001)　　=　　(1101)

—密文　　解密钥匙　　明文

—(0100) XOR (1101)　　=　　(1101)

—密文　　明文　　解密密钥=加密密钥

通用密钥加密

- 通用密钥加密=替换+置换+XOR运算
  - —古典密码
  - —流密码
  - —对称密钥加密算法
- 通用密钥加解密应用的密钥管理是个难题
  - —密钥个数=$C_n^2 = n*(n-1)/2$
  - —100个人相互通信所需密钥个数：4950
- 密钥一旦泄漏，则数据机密性被破坏
- 运算性能高，适合于大数据量加密

流密码

- 按比特位逐次进行加密和解密
  - —移动通信加密常用
  - —实时加密和实时解密
- RC4和SEAL是具有代表性的流密码
- 密钥：“伪”随机序列

明文　1 1 0 0 1 …… ……

XOR →

密钥　1 0 1 1 0 …… ……

以1bit为单位顺序进行加密

密文　1

流密码特点

- 运算速度快
  - —适用于实时加解密应用
  - —适合硬件化实现
- “伪”随机序列的随机性和长度是加密强度的重要保证
  - —理论上，如果密钥是真随机序列且密钥长度大于明文长度，则绝对安全
  - —现实中，很难实现真随机和无限长
    - – 计算机无法实现真随机

流密码安全性

- 密钥流
  - —周期：T
    - 如果存在一个固定的T，使得密钥流每隔T个符号后就出现重复，则称该流密码是周期的
  - —安全性需求
    - 极大的周期
    - 良好的统计特性
    - 线性不可预测性要充分大

流密码分类

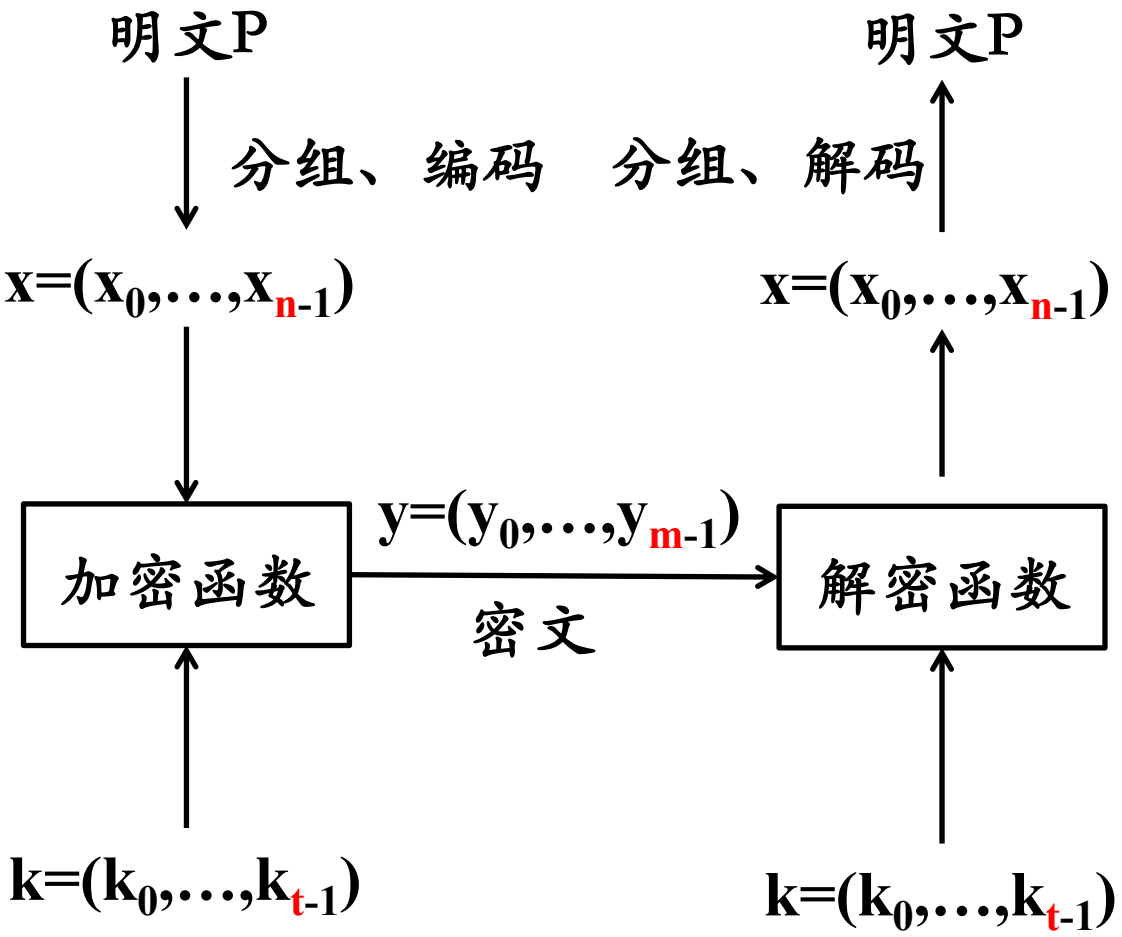
- 同步流密码
  - —密钥流与明文符号无关
  - —密钥流生成器+加/解密变换器
  - —加密密钥流和解密密钥流需要同步机制
  - —流密码的主流工作方式
- 自同步流密码
  - —密钥流与明文符号有关

本章内容提要

- 密码学简史
- 密码学基本概念
- 流密码
- 分组密码

分组密码框架

- m<n
  - —有数据压缩的分组密码
- m>n
  - —有数据扩展的分组密码
- m=n
  - —既无数据扩展也无数据压缩的分组密码



典型分组密码

- DES
  - —数据加密标准
- AES
  - —高级加密标准，DES的接班者
- IDEA
  - —国际数据加密算法
- SMS4
  - —国家密码管理局于2006年1月公布
  - —我国目前公布的第一个也是唯一一个密码算法

分组密码的设计原则——针对安全性需求

- 扩散
  - —将明文的统计特性散布到密文中去
  - —实现方式：明文变化一位，密文变化多位
- 混淆
  - —使密文和密钥之间的统计关系变得尽可能复杂
  - —使攻击者即使掌握了密文的统计特性，也无法推测密钥

分组密码的设计原则——针对实现需求

- 分块
  - —可并行化
- 使用简单运算
  - —软件实现：加法、乘法、移位
  - —硬件实现：加密与解密可用同样的器件来实现，且尽量使用规则结构

Feistel密码结构

- DES的设计基础之一
- 乘积密码
  - —顺序执行两个或多个基本密码系统，使得最后的密码强度高于每个基本密码系统产生的结果
  - —大量使用异或运算

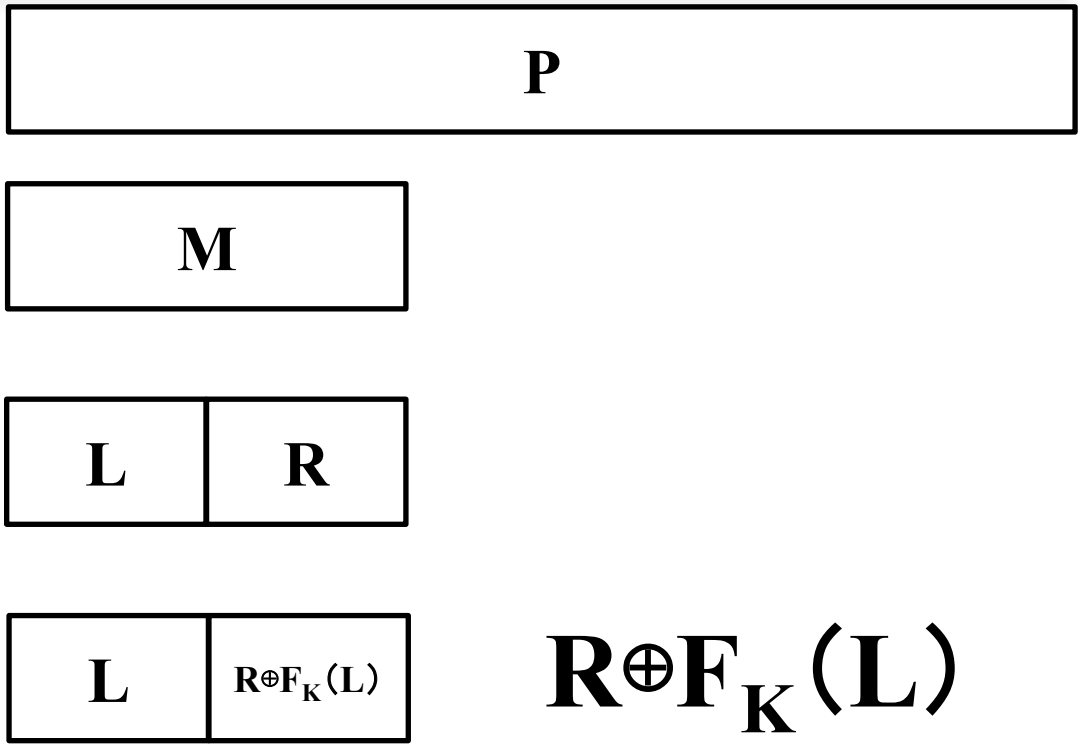
Feistel基本加密算法

❶ P分组得到M

❷ M分块得到L和R

❸ F_K变换

- ❶ 右F变换：RF_K
- ❷ 左F变换：LF_K



Feistel密码结构

- 轮函数f
  - —$f(R_{i-1},K_i)$
  - —（每轮基本加密运算的）子密钥K_i
- 在进行完n轮迭代后，左、右两部分再合并到一起以产生密文分组
  - —其中第i轮迭代的输入为前一轮的输出的函数

$$L_i=R_{i-1}$$
$$R_i=L_{i-1}\oplus f(R_{i-1},K_i)$$

Feistel网络特性

- 分组大小
- 密钥长度
- 轮数
- 子密钥产生算法
- 轮函数

安全性越高，加密速度越慢

Feistel解密

- 算法过程和加密算法过程完全相同
  - —加密与解密采用统一算法
  - —有利于硬件实现
- 输入密文，输出明文
- 使用子密钥次序与加密过程相反

DES背景(1/3)

- 发明人
  - —美国IBM公司 W. Tuchman 和 C. Meyer
  - —1971-1972年研制成功
- 基础
  - —1967年美国Horst Feistel提出的理论
- 产生
  - —美国国家标准局（NBS)1973年5月到1974年8月两次发布通告，公开征求用于电子计算机的加密算法。经评选从一大批算法中采纳了IBM的LUCIFER方案

DES背景(2/3)

- 标准化
  - —DES算法1975年3月公开发表
  - —1977年1月15日由美国国家标准局颁布为数据加密标准（Data Encryption Standard）
  - —1977年7月15日生效

DES背景(3/3)

- 美国国家安全局（NSA, National Security Agency)参与了美国国家标准局制定数据加密标准的过程。NBS接受了NSA的某些建议，对算法做了修改，并将密钥长度从LUCIFER方案中的128位压缩到56位
- 1979年，美国银行协会批准使用DES
- 1980年，DES成为美国标准化协会(ANSI)标准
- 1984年2月，ISO成立的数据加密技术委员会(SC20)在DES基础上制定数据加密的国际标准工作

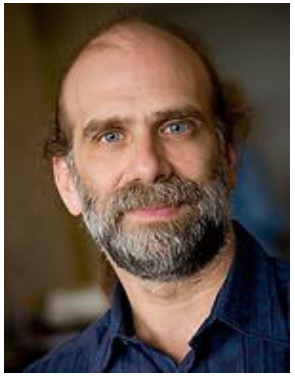
对称密码应用

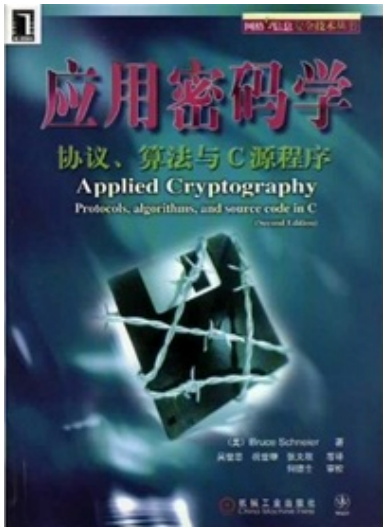
加密工具分享（1/2)

- Truecrypt
- Windows的文件系统加密

加密工具分享（2/2）

- Password Safe
  - —基于Twofish加密算法的开源密码管理工具
  - —Bruce Schneier，世界知名密码学专家
  - —AES候选算法：Blowfish、Twofish的联合作者
  - —著有密码学经典教材





课后思考题

- 试分析保密通信系统模型中所有可能的风险点？