정보보호론

암호이론과 보안개요

한림대학교 소프트웨어융합대학 조효진

Contents

- 암호학 소개
- 수학적 배경지식
- 고전암호
- 암호시스템의 안전성
- 정보보호 서비스

암호학 소개

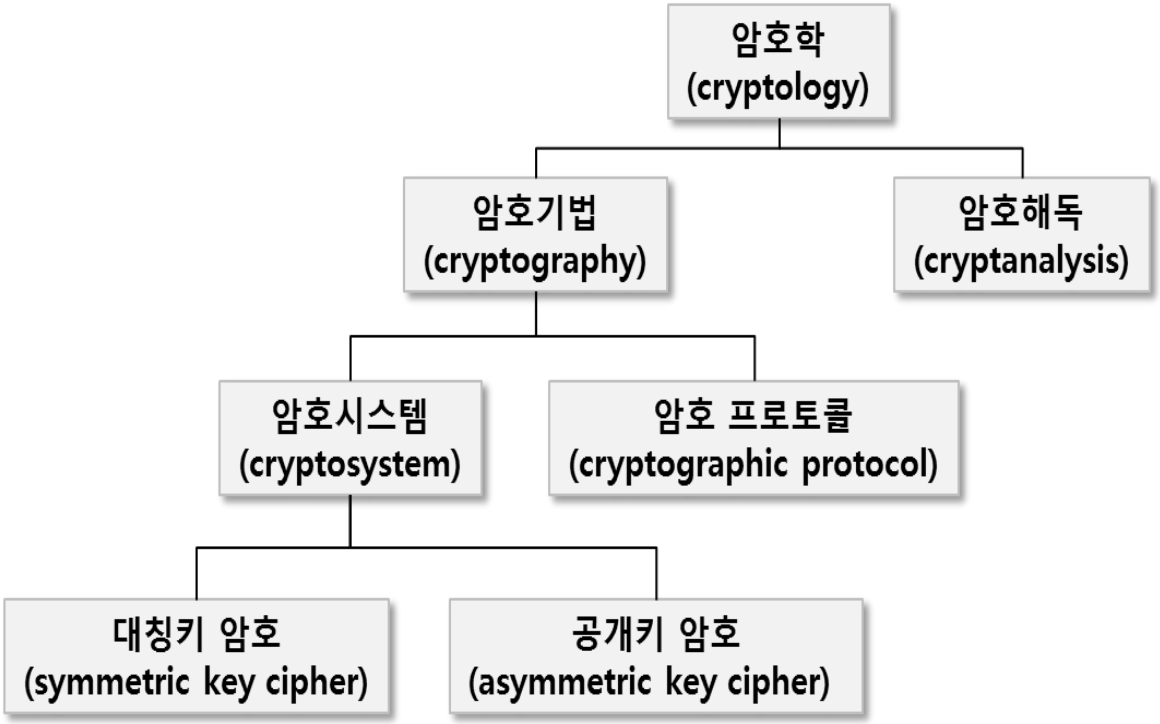
❒ 암호학(cryptology)

- 암호학은 보안시스템의 가장 중요한 부분이기도 하지만 그 자체로는 쓸모가 없다.
  - 웹(web)의 취약점을 찾는 공격자는 암호를 공격하지 않고도 버퍼 오버플로우(buffer overflow) 등을 이용하여 공격
  - 즉 "A security system is only as strong as its weakest link"

- 하지만, 다양한 암호 프리미티브를 통해 공격자의 공격행위를 제한하는 일은 필수적으로 진행되어야 함
  - Security by Design

암호학 소개

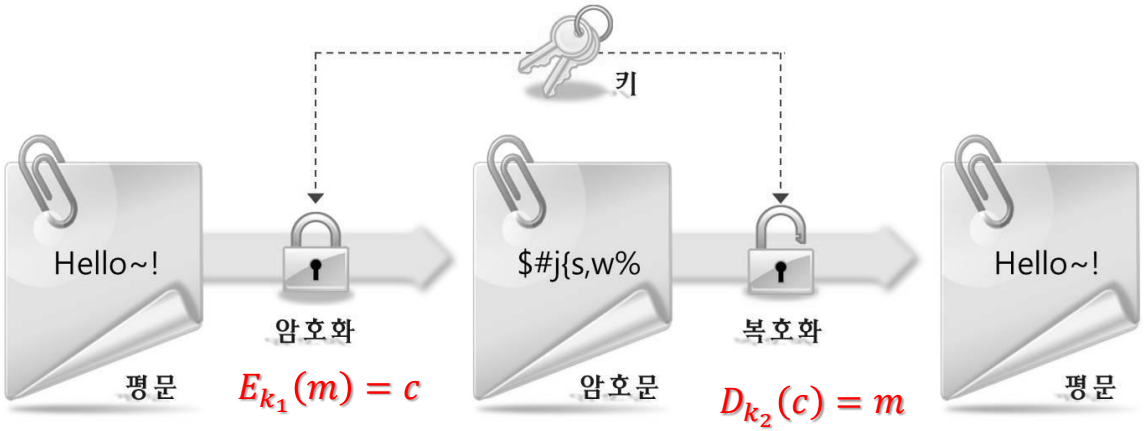
❒ 암호학(cryptology)의 분류



암호학 소개

❒ 암호기법(cryptography)

- 그리스어로 "비밀(secret)" 을 의미하는 kryptos와 "쓰다(write)" 를 의미하는 gráphō의 합성어
- 즉 메시지의 기밀성(confidentiality)을 제공하기 위하여 사용. 현재는 메시지를 공격자로부터 안전하게 보호하기 위하여 메시지를 변화하는 과학이나 기술을 의미.
- Key: uniformly distributed random string
- Symmetric (대칭키) if k1 = k2, Otherwise, asymmetric (공개키)



Correctness: $D_{k_2}(E_{k_1}(m)) = m$

Appendix #1

❒ 암호학 관련 대표 용어 정리

- Plaintext (평문): 암호화되기 전의 메시지를 의미함
- Encryption algorithm (암호화 알고리즘)
- Secret key (암호화 키): 암호 알고리즘의 안정성은 암호화 키에 의존함
- Ciphertext (암호문): 암호화 된 후의 메시지를 의미함
- Decryption algorithm (복호화 알고리즘)

Appendix #1

If P is the plaintext, C is the ciphertext, and K is the key,

Encryption: $C = E_k(P)$ Decryption: $P = D_k(C)$

In which, $D_k(E_k(x)) = E_k(D_k(x)) = x$

We assume that Bob creates $\mathbf{P_1}$***; we prove that*** $\mathbf{P_1 = P}$**:**

Alice: $C = E_k(P)$

Bob: $P_1 = D_k(C) = D_k(E_k(P)) = P$

암호학 소개

❒ Symmetric vs Public Key Algorithms

- Secure communication using symmetric key k.
  - Alice → Bob : $E_k(m) = c$
  - Bob decrypts c by using k
- How can Alice and Bob share k?
  - Public Key Algorithms solves this problem!
  - Encrypt k using Bob's public key k_{pub} , i.e., $E_{pub_k}(k)$. Bob then decrypts $E_{pub_k}(k)$ using his private key k_{priv} to obtain k.
  - 공개키 암호 파트에서 자세히 배움!

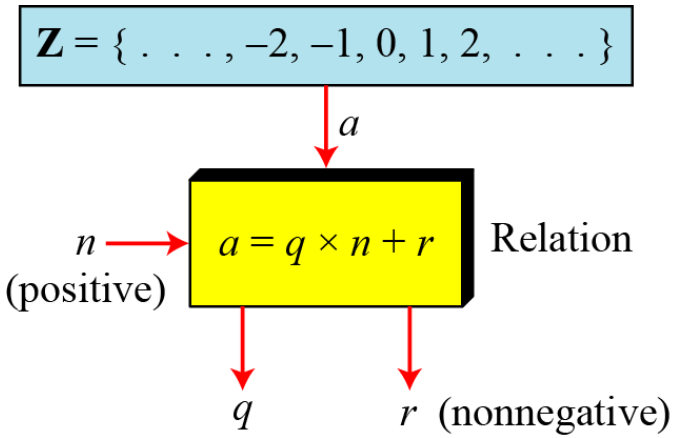
암호학 소개: Kerckhoff's Principle

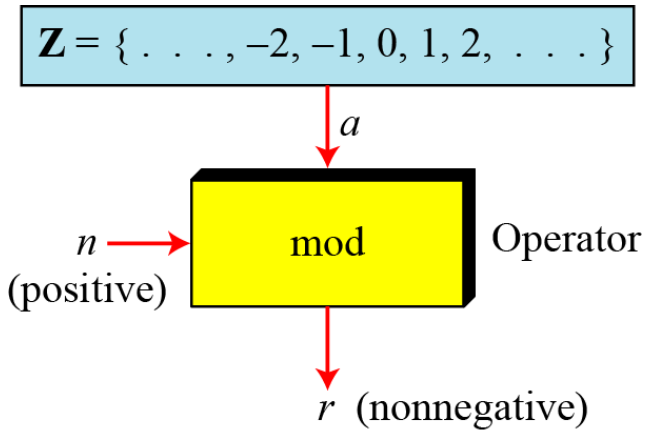
- Kerckhoffs의 원리: 암호 알고리즘은 알고리즘의 모든 내용이 공개되어도 키가 노출되지 않으면 안전해야 한다.
  - 짧은 길이의 키를 안전하게 보관하는 것은 키 보다 수천배의 사이즈인 암호 알고리즘 전체를 안전하게 보관하는 것 보다 용이. 또한 암호시스템은 역공학 등으로 노출될 수 있음
  - 키가 노출되었을 때 키를 변경하는 것이 새로운 암호시스템을 설계하는 것보다 훨씬 용이
  - 암호시스템은 보통 다수의 사용자를 위하여 운영되며, 모든 사용자는 동일한 암호 알고리즘을 사용. 이 경우 암호 통신을 하는 당사자들마다 상이한 암호시스템을 사용하는 것 보다는 동일한 암호시스템을 사용하면서 키만 다르게 설정하는 것이 실용적. → 표준화

수학적 배경지식

❒ 모듈라 연산(modular arithmetic)

- 임의의 정수 a를 양의 정수 n으로 나누면 몫이 q가되고 음이 아닌 나머지 r을 얻는다.
  - $a = qn + r \quad 0 \leq r < n$
  - $23 = 4 \times 5 + 3$
  - $-17 = (-3) \times 5 + (-2) = (-4) \times 5 + 3$
- mod 연산
  - $a \bmod n = r$
  - $23 \bmod 5 = 3; \; -17 \bmod 5 = 3$





수학적 배경지식

❒ 모듈라 연산(modular arithmetic)

- mod 연산은 임의의 정수 a를 양의 정수 n으로 나누면 몫이 q가 되고 음이 아닌 나머지 r을 얻는다.

$$a = qn + r \quad 0 \le r < n$$

- mod 연산은 정수집합 Z_n (음의 정수 제외) 을 만듦

$$\mathbf{Z}_n = \{ 0, 1, 2, 3, \ldots, (n-1) \}$$

$$\mathbf{Z}_2 = \{ 0, 1 \} \qquad \mathbf{Z}_6 = \{ 0, 1, 2, 3, 4, 5 \} \qquad \mathbf{Z}_{11} = \{ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 \}$$

- 합동(Congruence)

$$2 \equiv 12 \pmod{10} \qquad 13 \equiv 23 \pmod{10}$$
$$3 \equiv 8 \pmod{5} \qquad 8 \equiv 13 \pmod{5}$$

수학적 배경지식

❒ 역원 (Inverses): 덧셈상의 역원, 곱셈상의 역원

- Z_n 상에서 덧셈상의 역원

$$a + b \equiv 0 \pmod{n}$$

- Z_n 상에서 곱셈상의 역원

$$a \times b \equiv 1 \pmod{n}$$

– In modular arithmetic, an integer may or may not have a multiplicative inverse. Number a has the mult. Inverse iff gcd(n,a) ≡ 1 (mod n)

고전암호

- 2 가지 원칙: 치환(Substitution)과 전치(Transposition)

- 암호 단위
  - 고전 암호 - 문자
  - 현대 암호 - 비트

- 공격 유형
  - 전사적 공격(Brute Force Attack) - 전수 키 탐색 공격(Exhaustive Key Search Attack), 현대 암호는 키의 길이가 길기 때문에 전사적 공격은 사실상 불가능
  - 빈도수 분석(Frequency Analysis) - 평문의 통계학적 특성이 암호문에 나타나는 성질을 이용하여 공격하는 방법
  - 암호 공격 (Cryptanalytic attack)
    - the nature of the algorithm plus perhaps some knowledge of the general characteristics of the plaintext or even some sample plaintext-ciphertext pairs.

고전암호: 치환 암호(Substitution Cipher)

❒ 치환 암호

- 단일 문자 치환 암호(Monoalphabetic Substitution Cipher): 평문의 한 문자와 암호문의 한 문자는 언제나 일대일 관계
- 다중 문자 치환 암호(Polyalphabetic Substitution Cipher): 평문의 한 문자와 암호문의 한 문자는 언제나 일대일 관계가 아님
- Plaintext and ciphertext in Z_{26}

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

고전암호: 치환 암호(Substitution Cipher)

❒ 단일 문자 치환 암호: 덧셈 암호(Additive Cipher)

- 시저 암호 (Caesar cipher)
  - E.g., 평문의 한 문자가 오른쪽 세 자리 뒤에 위치한 문자로 치환
- 암호화 : $c \equiv m + 3 \ (mod\ 26),\ m \in \mathbb{Z}_{26}$
- 복호화 : $m \equiv c - 3 \ (mod\ 26),\ c \in \mathbb{Z}_{26}$

m	A	B	C	D	E	F	G	H	I	J	K	L	M
	0	1	2	3	4	5	6	7	8	9	10	11	12
$c \equiv m + 3$ (mod 26)	d	e	f	g	h	i	j	k	l	m	n	o	p
	3	4	5	6	7	8	9	10	11	12	13	14	15
m	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	13	14	15	16	17	18	19	20	21	22	23	24	25
$c \equiv m + 3$ (mod 26)	q	r	s	t	u	v	w	x	y	z	a	b	c
	16	17	18	19	20	21	22	23	24	25	0	1	2

고전암호: 치환 암호(Substitution Cipher)

단일 문자 치환 암호: 덧셈 암호(Additive Cipher)

- 암호화 : $c \equiv m + k \ (mod\ 26),\ m \in \mathbb{Z}_{26}$
- 복호화 : $m \equiv c - k \ (mod\ 26),\ c \in \mathbb{Z}_{26}$
- 덧셈암호에 대한 전사적 공격 (Brute force attack)
  - 가능한 키의 경우의 수는? 26개

고전암호: 치환 암호(Substitution Cipher)

❒ 일반적인 치환암호

- 각 문자에 적용되는 k가 다를 수 있음

평문	A	B	C	D	E	F	G	H	I	J	K	L	M
암호문	l	z	h	t	g	s	b	c	v	d	y	n	m
평문	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
암호문	x	j	i	k	u	a	w	f	o	r	q	e	p

- 치환암호에 대한 전사적 공격이 가능한가?
  - 가능한 키의 개수는 총 $26 \times 25 \times \cdots \times 1 = 26!$ 개
  - 전사적 공격을 이용하여 공격자가 키를 찾는 것은 불가능
    - 하나의 키를 시도하여 복호화하는데 소요되는 시간이 1초라 가정하면, 전사적 공격으로 모든 키를 이용하여 복호화하는데 소요되는 시간은 403,291,461,126,605,635,584,999,999초

고전암호: 치환 암호(Substitution Cipher)

- **하지만, 치환암호의 경우도 안전하지 않음**
  - 빈도수 공격 : 통계적 특성 이용
    - 988,968 개의 영어 단어 중, "E" 가 사용되는 횟수는 12.7%

Letter	Probability	Letter	Probability	Letter	Probability
A	0.082	B	0.015	C	0.028
D	0.043	E	0.127	F	0.022
G	0.020	H	0.061	I	0.070
J	0.002	K	0.008	L	0.040
M	0.024	N	0.067	O	0.075
P	0.019	Q	0.001	R	0.060
S	0.063	T	0.091	U	0.028
V	0.010	W	0.023	X	0.001
Y	0.020	Z	0.001		

고전암호: 치환 암호(Substitution Cipher)

- **하지만, 치환암호의 경우도 안전하지 않음**
  - 치환 암호를 이용하여 어떠한 메시지를 암호화하였다고 하더라도, 평문의 한 문자와 암호문의 한 문자가 일대일 대응관계이기 때문에 이러한 통계학적인 특성은 그대로 유지되게 됨
  - 예를 들어, 암호문에서 가장 많이 사용된 문자는 "E" 라고 짐작할 수 있음
  - 또한, 만약 "THE__E" 에서 한 문자만을 해독하지 못했다고 하면 이 단어는 "THERE" 일 것이라고 추측할 수 있게 됨
  - 위의 표에 나타난 빈도수와 같이 한 문자 이외에 연속해서 사용되는 두 문자, 세 문자에 대한 빈도도 유용하게 사용될 수 있음
    - TH, HE, IN, ER, AN, RE, ED, ON, ES, ST, EN, AT, TO, NT, HA, ND, OU, EA, NG, AS, OR, TI, IS, ET, IT, AR, TE, SE, HI, OF
    - THE, ING, AND, HER, ERE, ENT, THA, NTH, WAS, ETH, FOR, DTH

고전암호: 치환 암호(Substitution Cipher)

❒ 다중 문자 치환 암호: 비제네르 암호 (Vigenère Cipher)

- 길이가 d인 키워드를 암호화 키로 사용
- $K = k_1 k_2 k_3 \ldots k_d,\ f_i(m) = (m + k_i) \bmod n$

평문	T	H	I	S	I	S	A	S	E	C	R	E	T	M	E	S	S	A	G	E
	19	7	8	18	8	18	0	18	4	2	17	4	19	12	4	18	18	0	6	4
키	C	I	P	H	E	R	C	I	P	H	E	R	C	I	P	H	E	R	C	I
	2	8	15	7	4	17	2	8	15	7	4	17	2	8	15	7	4	17	2	8
암호문	V	P	X	Z	M	J	C	A	T	J	V	V	V	U	T	Z	W	R	I	M
	21	15	23	25	12	9	2	0	19	9	21	21	21	20	19	25	22	17	8	12

주기 d = 6

고전암호: 치환 암호(Substitution Cipher)

❒ 다중 문자 치환 암호: 비제네르 암호 (Vigenère Cipher)

- 공격: KASISKI METHOD
- period "d"를 결정하는 방법

LIOMWGFEGGDVWGHHCQUCRHRWAGWIOWQLKGZETKKMEVLWPCZVGTH-
VTSGXQOVGCSVETQLTJSUMVWVEUVLXEWSLGFZMVVWLGYHCUSWXQH-
KVGSHEEVFLCFDGVSUMPHKIRZDMPHHBVWVWJWIXGFWLTSHGJOUEEHH-
VUCFVGOWICQLTJSUXGLW

String	*First Index*	*Second Index*	*Difference*
JSU	68	168	100
SUM	69	117	48
VWV	72	132	60
MPH	119	127	8

고전암호: 치환 암호(Substitution Cipher)

❐ 다중 문자 치환 암호: 비제네르 암호 (Vigenère Cipher)

- 차이의 GCD는 4 → 키 길이는 4의 배수
- First try l = 4. (C1, C2, C3, C4에 빈도수 분석)

LIOMWGFEGGDVWGHHCQUCRHRWAGWIOWQLKGZETKKMEVLWPCZVGTH-
VTSGXQOVGCSVETQLTJSUMVWVEUVLXEWSLGFZMVVWLGYHCUSWXQH-
KVGSHEEVFLCFDGVSUMPHKIRZDMPHHBVWVWJWIXGFWLTSHGJOUEEHH-
VUCFVGOWICQLTJSUXGLW

C1: LWGWCRAOKTEPGTQCTJVUEGVGUQGECVPRPVJGTJEUGCJG

P1: *jueuapymircneroarhtsthihytrahcieixsthcarrehe*

C2: IGGGQHGWGKVCTSOSQSWVWFVYSHSVFSHZHWWFSOHCOQSL

P2: *usssctsiswhofeaeceihcetesoecatnpntherhctecex*

C3: OFDHURWQZKLZHGVVLUVLSZWHWKHFDUKDHVIWHUHFWLUW

P3: *lcaerotnwhiwedssirsiirhketehretltiideatrairt*

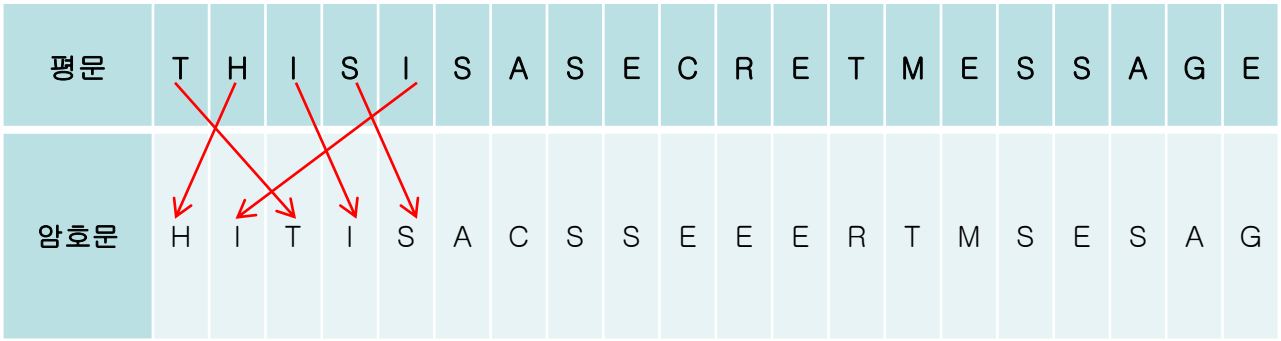
C4: MEVHCWILEMWVVXGETMEXLMLCXVELGMIMBWXLGEVVITX

P4: *iardysehaisrrtcapiafpwtethecarhaesfterectpt*

고전암호: 전치 암호(Transposition Cipher)

❒ 평문 메시지의 문자들을 재배열

- 전치 암호의 암호화 함수를 π 라 하고 문자열의 길이를 l이라 하면,
  - $\pi = (\pi(1), \dots, \pi(l))$
- $\pi(i)$: 평문에서 i번째 위치에 있는 문자의 암호문에서의 위치
  - 예제 : $\pi = (\pi(1), \dots, \pi(5)) = (3,1,4,5,2)$



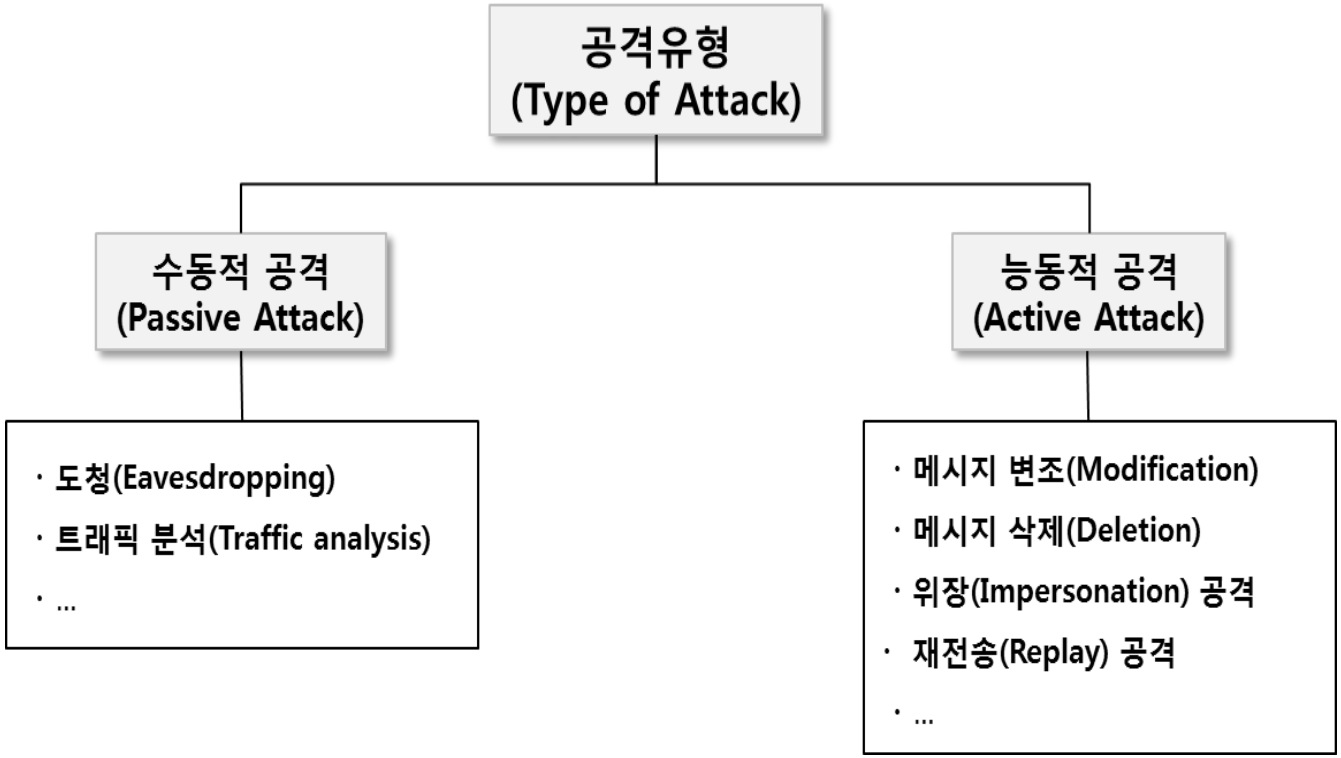
고전암호: 전치 암호(Transposition Cipher)

- 전치 암호도 전사적 공격과 통계 공격으로부터 안전하지 않음
  - 메시지의 길이가 30인 경우 → $1! + 2! + 3! + \cdots + 30!$
  - 블록이 30의 약수! → 30=1×2×3×5 → 30의 인수인 1, 2, 3, 5, 6, 10, 15, 30를 이용하여 $1! + 2! + 3! + 5! + 6! + 10! + 15! + 30!$

고전 암호는 쉽게 공격되므로, 암호 통신을 위해 새로운 암호가 필요하게 되었음
→ 대칭키 (e.g., AES), 공개키 (e.g., RSA)암호가 설계됨

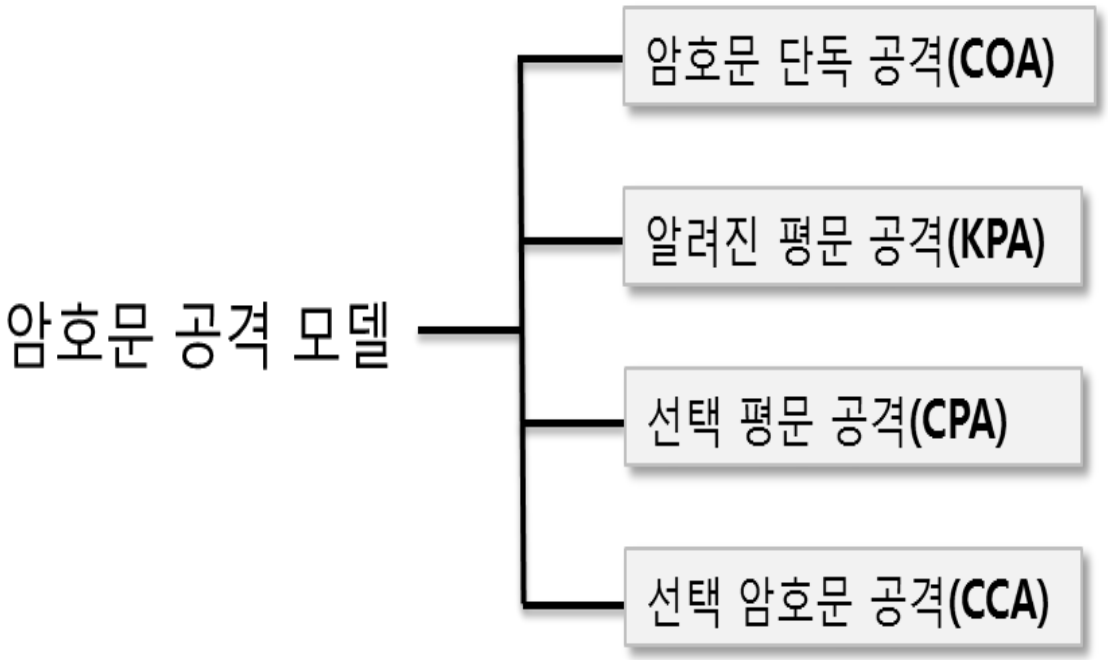
암호시스템의 안전성

❐ 공격유형(Type of Attack)



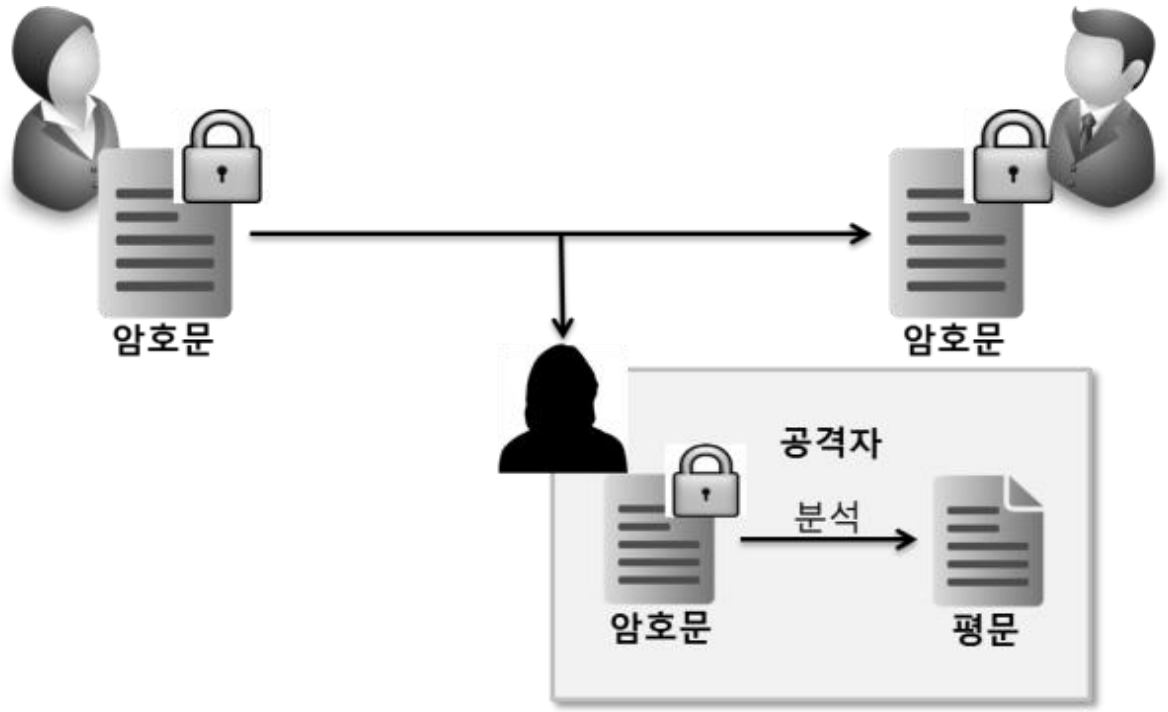
암호시스템의 안전성

❐ 공격모델(Attack Model)



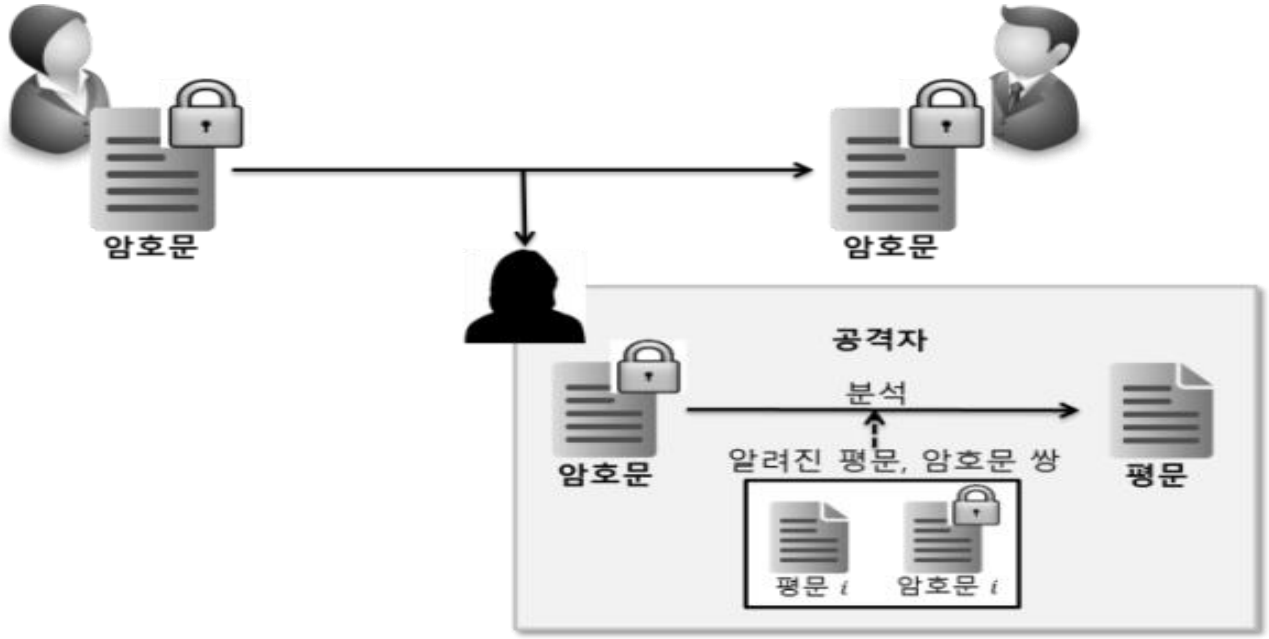
암호시스템의 안전성

- 공격모델: 암호문 단독 공격(Ciphertext Only Attack, COA)



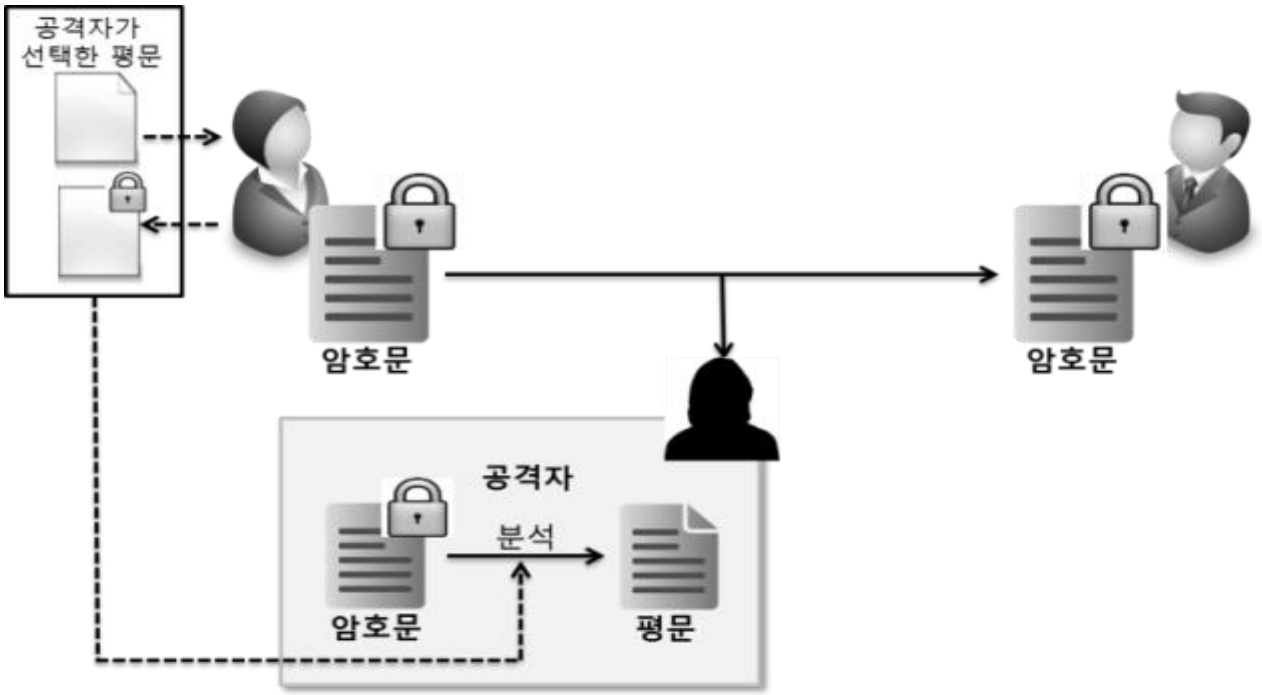
암호시스템의 안전성

- 공격모델: 알려진 평문 공격(Known Plaintext Attack, KPA)



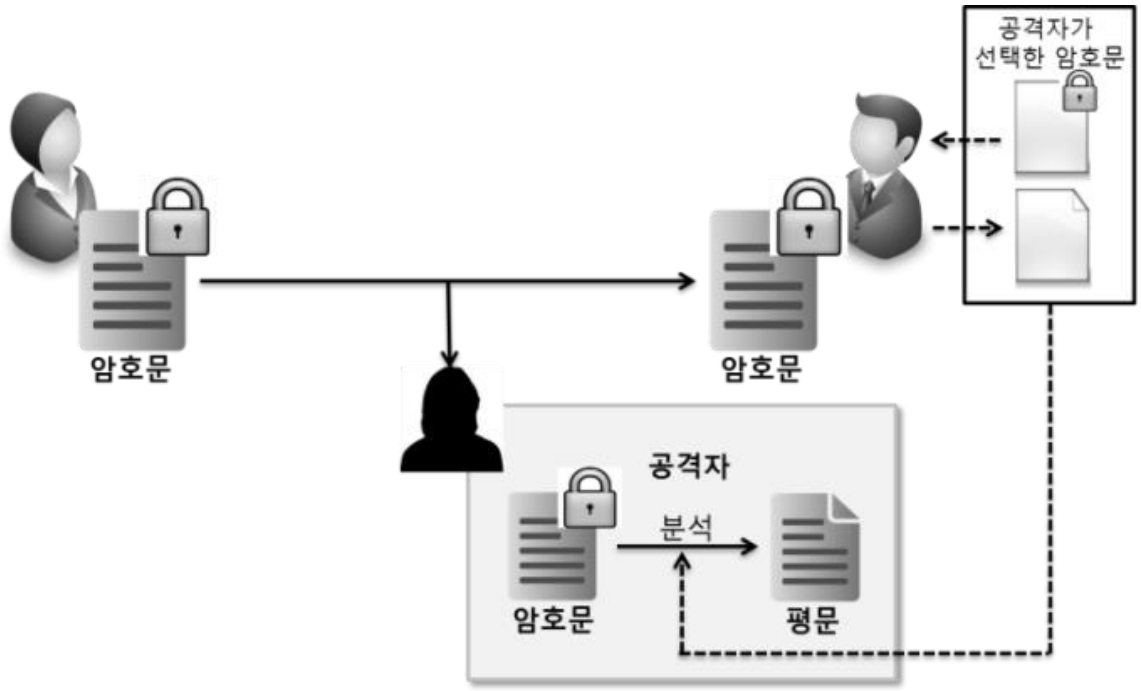
암호시스템의 안전성

- 공격모델: 선택 평문 공격(Chosen Plaintext Attack, CPA)



암호시스템의 안전성

- 공격모델: 선택 암호문 공격(Chosen Ciphertext Attack, CCA)



암호시스템의 안전성

❒ 안전한 암호란?

- 공격자가 암호시스템에서의 비밀키를 얻을 수 있다면, 그 암호시스템은 완전히 파괴되었다고 말할 수 있음
- 하지만 공격자가 키를 알아낼 수는 없지만 암호문의 일부분, 혹은 한 글자만 알 수 있는 경우, 안전하다고 할 수 있을까?
  - 예를 들어 한 암호문이 중요한 금액(연봉, 혹은 입찰가 등)인 경우, 공격자가 이 암호문의 금액이 1억원 이상인지 아닌지를 판단할 수 있다면 이 암호는 공격자가 필요로하는 중대한 정보를 노출시키고 있다고 볼 수 있음

정보보호 서비스

- 3대 정보보호 서비스(NIST)
  - 기밀성(Confidentiality)
  - 무결성(Integrity)
  - 가용성(Availability)

- 위의 정보보호 서비스 이외에 환경에 따라 요구되는 서비스는 다양하며 다음과 같다.
  - 인증(Authentication)
    - 개체 인증(Entity Authentication) : 개체가 정당한(혹은 개체가 주장하는) 개체인지를 확인하는 성질을 의미한다.
    - 메시지 인증(Message Authentication) : 수신된 메시지가 정당한 송신자로부터 전송된 것인지를 확인하는 성질을 의미한다. 즉 수신된 메시지의 송신자를 인증하는 과정이다.
  - 부인방지(Non-Repudiation)
  - 접근제어(Access Control)

Assignment

- 다음의 모듈라 연산을 하시오.
  - ① $34\ mod\ 7$
  - ② $45\ mod\ 5$
  - ③ $-12\ mod\ 5$
  - ④ $-27\ mod\ 4$

- 시저 암호를 이용하여 평문 “CAESAR CIPHER”를 암호화 하시오.
  - Hint: 시저 암호의 key는 3임

Assignment

- 비제네르 암호를 이용하여, 다음 평문을 암호화 하시오. (암호화 키: KEY)
  - 평문: THISISVERYIMPORTANT

- π=(2,4,1,5,6,3)일 때, 전치 암호를 이용하여 암호문 “VHNAEAEIYCDA”를 복호화 하시오.

Assignment

❒ 과제제출

- 6월 26일 정오 12시까지
- 늦은 제출 시, 감점
- 과제 copy 시, 관련된 과제들 모두 0점 처리
- 제출양식: hwp or pdf

Thank you