교과목 : 정보보호

2. Crypto Basics

2023학년도 2학기
Suk-Hwan Lee

References

Textbook

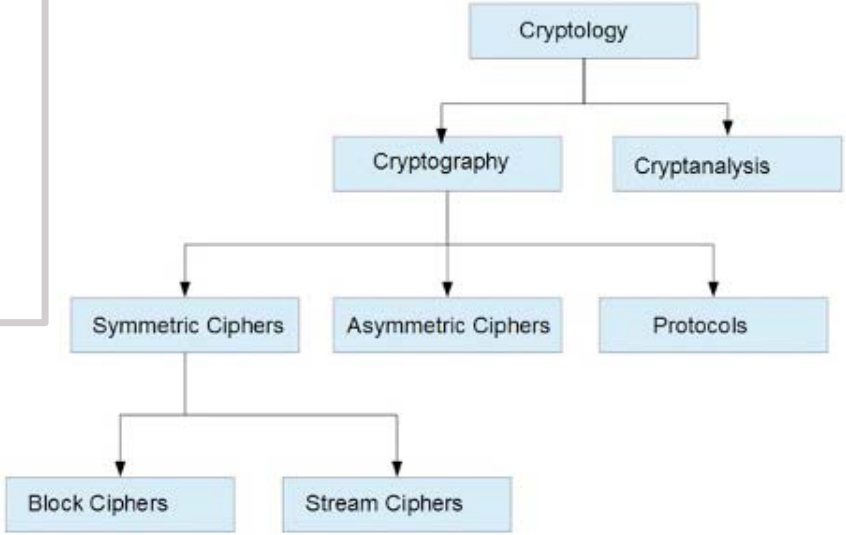
- Mark Stamp, Information Security: Principles and Practice, Second edition, & Lecture Note
- William Stallings, Cryptography and Network Security, Seventh Edition

참조

- Stanford Univ., https://crypto.stanford.edu/~dabo/courses/OnlineCrypto/
- 부산대, Computer Security, Lecture Note
- 단국대, Introduction to Software Security, Lecture Note
- 명지대, Computer Security, Lecture Note
- 서울과학기술대, Information Protection Theory, Lecture Note
- York Univ. Network Security & Forensics, Lecture Note
- 해시넷, http://www.hash.kr/
- K.M. Kareem, Univ. of Sulaimani, Cryptography with Classical Cipher, Lecture Note
- Wikipedia
- Cryptographics, https://cryptographics.info/all-cryptographics/#
- etc.....

1. Crypto

- **Cryptology** – The art and science of making and breaking "secret codes" (Cryptograph+Cryptanalysis)
- **Cryptography** – making "secret codes" : Study of encryption principles/algorithms
- **Cryptanalysis** – breaking "secret codes" : Study of principles/algorithms of deciphering ciphertext without knowing key
- **Crypto** – all of the above (and more)



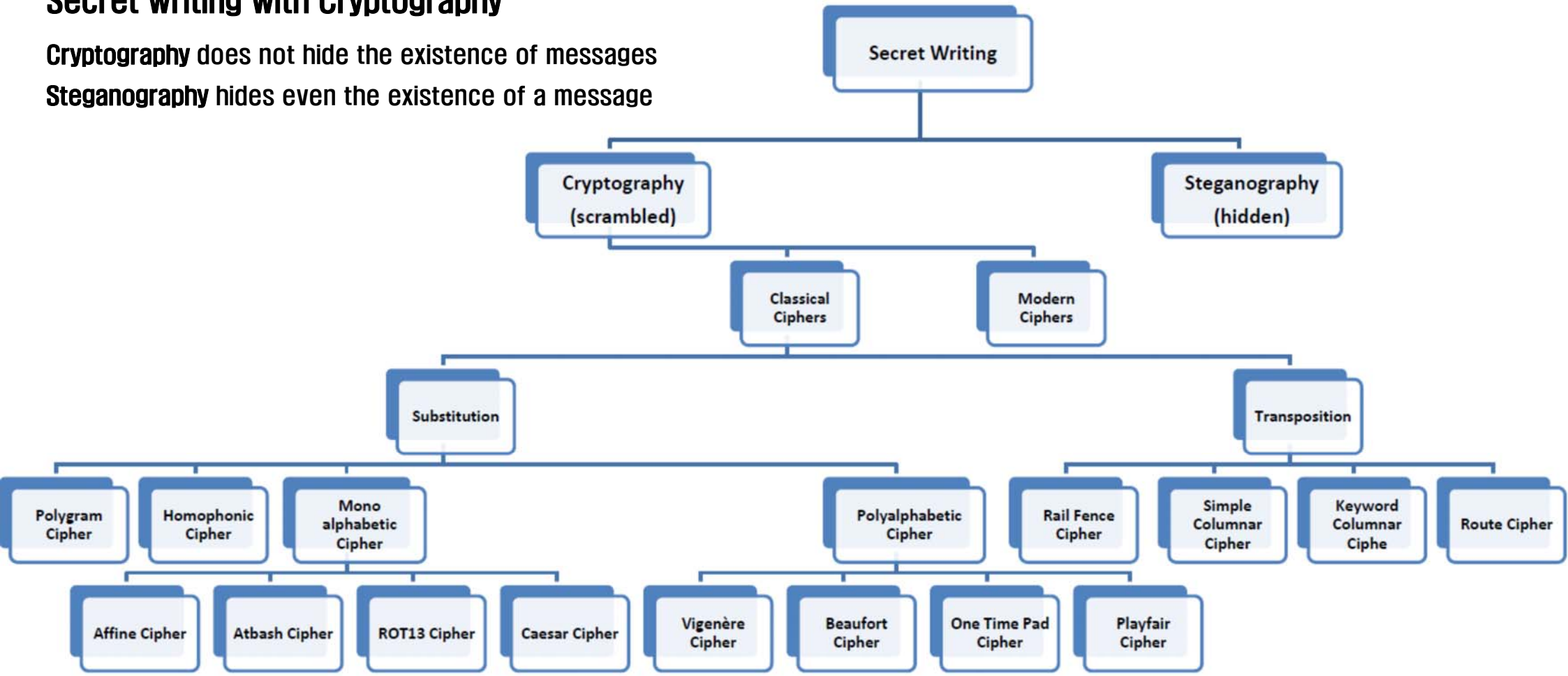
How to Speak *Crypto*

- A **cipher** or **cryptosystem** is used to **encrypt** the **plaintext**
- The result of encryption is **ciphertext**
- We **decrypt** ciphertext to recover plaintext
- A **key** is used to configure a cryptosystem
- A **symmetric key** cryptosystem uses the same key to encrypt as to decrypt
- A **public key** cryptosystem uses a **public key** to encrypt and a **private key** to decrypt (sign)

Secret writing with Cryptography

Cryptography does not hide the existence of messages

Steganography hides even the existence of a message



[그림 출처] K.M. Kareem, Univ. of Sulaimani, Cryptography with Classical Cipher, Lecture Note

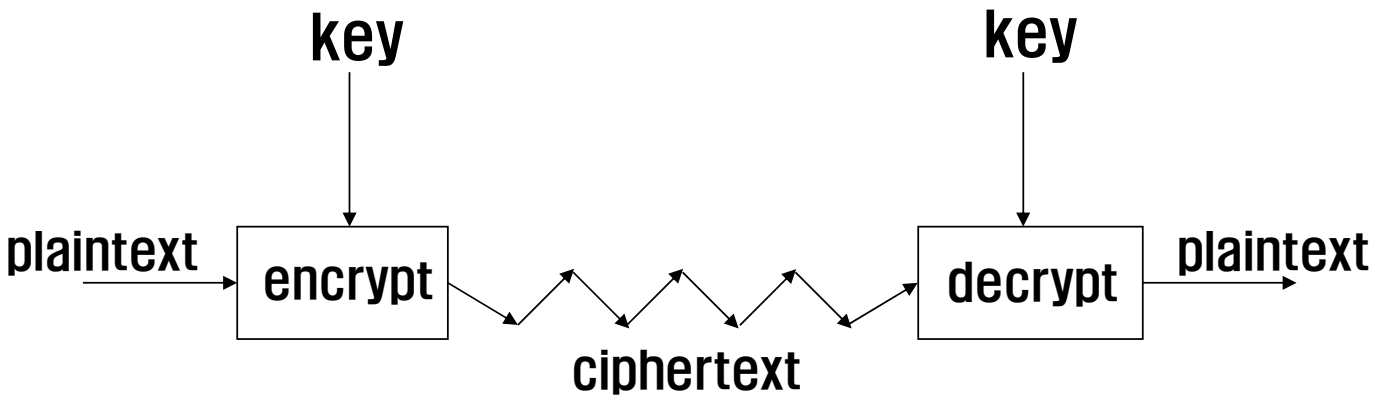
- Basis assumption
  - ✓ The system is completely known to the attacker (암호 알고리즘은 비밀이 아님)
  - ✓ Only the key is secret

- Also known as **Kerckhoffs Principle**
  - ✓ Crypto algorithms are not secret (암호 체계는 공격자가 키를 제외한 모든 내용 또는 알고리즘을 알더라도 안전해야 함)

- Why do we make this assumption?
  - ✓ Experience has shown that secret algorithms are weak when exposed
  - ✓ Secret algorithms never remain secret
  - ✓ Better to find weaknesses beforehand

Steganography – Unlimited algorithms
(알고리즘을 찾는 것이 매우 어려움)

Crypto as Black Box



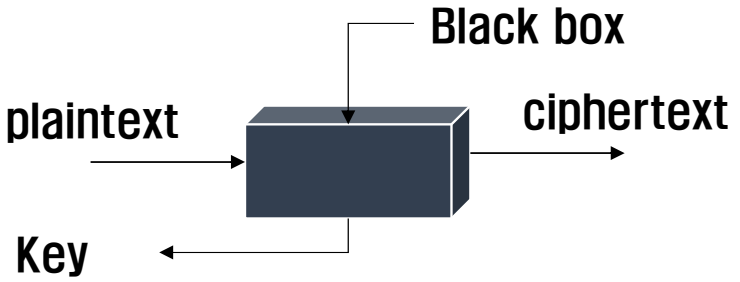
A generic use of crypto

- Plaintext : the original message
- Ciphertext : the encrypted (coded) message
- Cipher : algorithm for transforming plaintext to ciphertext

- Key : info used in cipher known only to sender/receiver
- Encipher(Encrypt) : Convert plaintext to ciphertext
- Decipher(Decrypt) : Recover plaintext from ciphertext

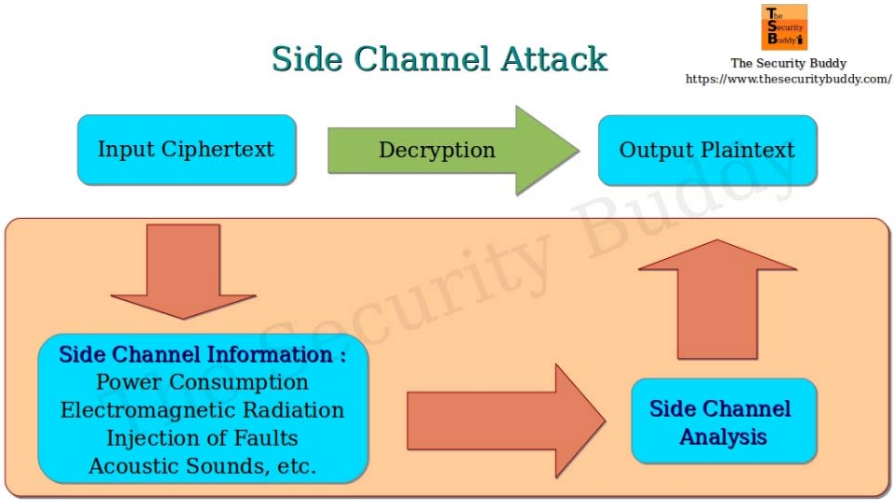
Attacks in Cryptography

- Black-box attack
  - ✓ Watches inputs and outputs
  - ✓ Controls input text
  - ✓ No visibility of execution
  - ✓ Tries to deduce the key from a list {(plaintext, ciphertext)}



❖ *Side-channel attack*

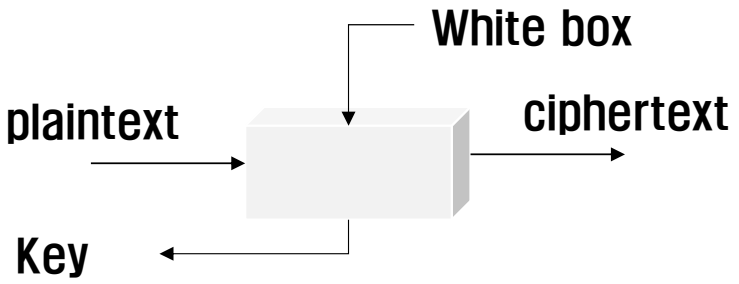
- Cryptosystem is implemented on physical devices. It is very difficult to break the cryptographic algorithm
- But usually, during cryptographic computation, the device reveal some information in terms of Executing time, Electromagnetic radiation, Power consumption (Side-channel information)
- An attack uses this side-channel information to determine the secret keys and break the cryptosystem



[그림 출처] https://www.thesecuritybuddy.com/vulnerabilities/what-is-side-channel-attack/

Attacks in Cryptography

- White-box attack
  - ✓ Attacker knows algorithm and can observe everything
  - ✓ Watches inputs, outputs, and intermediate calculations
  - ✓ Controls input text
  - ✓ Full visibility into Memory (debuggers and emulators)
  - ✓ Target for attack : Implementation of cryptography, Secret key

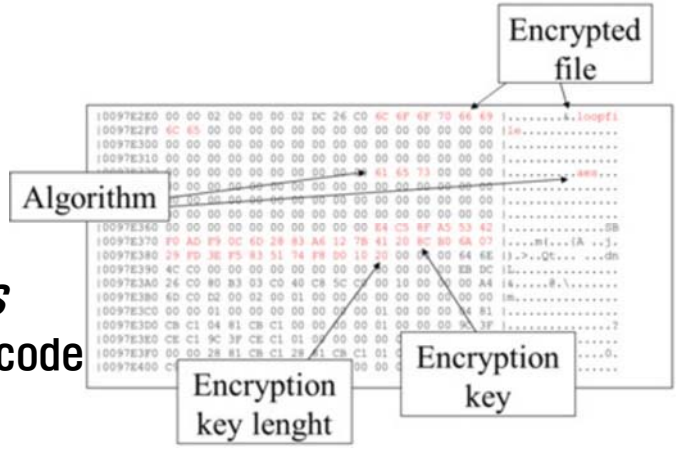


❖ *Key Whitening Attack*
- Zero lookup tables (Such as S-box) using hex editor
- Getting output of the penultimate (끝에서 2번째) operation
- Original AES key easily be derived

❖ *Format Analysis*
- Analyze binary code

❖ Entropy attack

❖ Code boot attack



1. Crypto

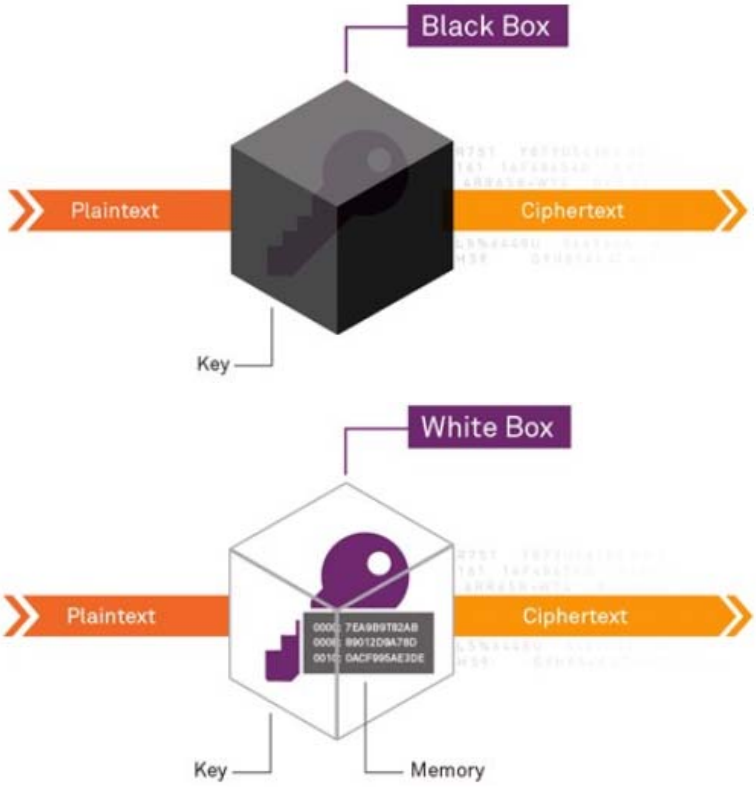
Attacks in Cryptography

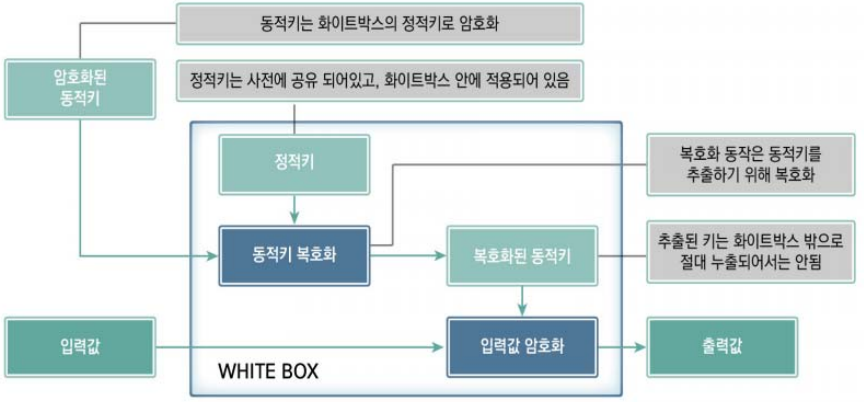
- Black-box cryptography (예전 암호화 기술)
  - ✓ 암호화 과정에는 암호화 키가 필요한데 이 암호화 키는 블랙박스로 가정한 암호화 장치 내부에 들어 있음
  - ✓ 크래커가 이 암화화 장치 내부는 들여다 볼 수 없다고 가정하나 2개의 입출력 값을 계속 관찰하여 어떤 패턴을 알아 낼 수 있음 (black-box attack)

    Y = algorithm1 (x, key1)

- White-box cryptography (WBC)
  - ✓ 공격자가 암호화 키를 유추할 수 없도록 하는 기술
  - ✓ 암호화 키가 데이터와 암호 알고리즘 속에 섞여 있어서 공격자가 암호키를 쉽게 볼 수 없음

    Y = algorithm2 (x)





Inside Secure사의 White-box 모듈 동작 구조

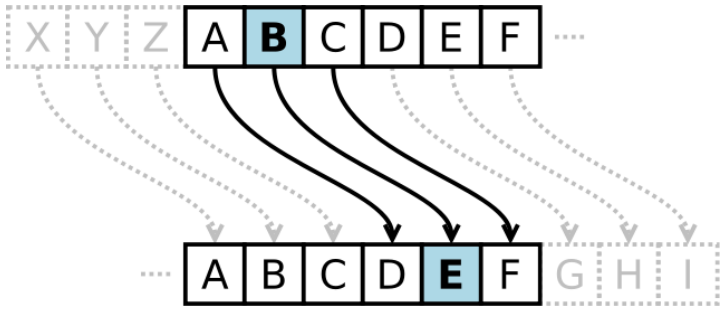
[출처] http://wiki.hash.kr/index.php/블랙박스_암호

Simple Substitution

- Plaintext: fourscoreandsevenyearsago
- Key:

Plaintext	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
Ciphertext	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

- Ciphertext: IRXUVFRUHDAGVHYHABHDUVDIR
- Shift by 3 is "Caesar's cipher"



Caesar's Cipher Decryption

- Suppose we know a Caesar's cipher is being used
- Ciphertext: VSRQJHEREVTXDUHSDQWU

Plaintext	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
Ciphertext	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

- Plaintext: spongebobsquarepants

Not-so-Simple Substitution

- Shift by n for some $n \in \{0,1,2,\cdots,25\}$
- Then key is n
- Example: key = 7

Plaintext	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
Ciphertext	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

Cryptanalysis I: Try Them All

- Given
  - ✓ A simple substitution (shift by n) is used
  - ✓ But the key is unknown
  - ✓ Given ciphertext: meqefscerhcsyeviekmvp
- How to find the key?
- Exhaustive key search
  - ✓ Only 26 possible keys — try them all!
  - ✓ Solution: key = 4 IAMABOYANDYOUAREAGIRL

Even-less-Simple Substitution

- Key is some permutation of letters
- Need not be a shift
- For example

Plaintext	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
Ciphertext	J	I	C	A	X	S	E	Y	V	D	K	W	B	Q	T	Z	R	H	F	M	P	N	U	L	G	O

- Then 26! > 2^{88} possible keys!
- Dominates the art of secret writing throughout the first millennium

Cryptanalysis II: Be Clever

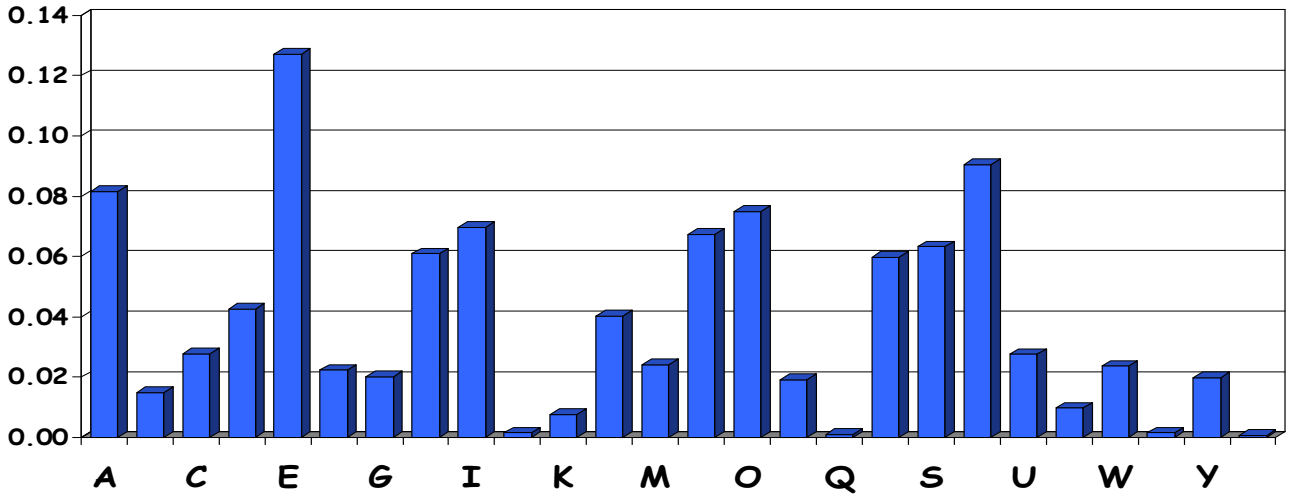
- We know that a simple substitution is used
- But not necessarily a shift by *n*
- Can we find the key given ciphertext:

PBFPVYFBQXZTYFPBFEQJHDXXQVAPTPQJKTOYQWIPBVWLXTOXBTFXQWAXBVCXQWAXFQJVWLEQNTOZQGG
QLFXQWAKVWLXQWAEBIPBFXFQVXGTVJVWLBTPQWAEBFPBFHCVLXBQUFEVWLXGDPEQVPQGVPPBFTIXPF
HXZHVFAGFOTHFEFBQUFTDHZBQPOTHXTYFTODXQHFTDPTOGHFQPBQWAQJJTODXQHFOQPWTBDHHIXQVAPB
FZQHCFWPFHPBFIPBQWKFABVYYDZBOTHPBQPQJTQOTOGHFQAPBFEQJHDXXQVAVXEBQPEFZBVFOJIWFFACF
CCFHQWAUVWFLQHGFXVAFXQHFUFHILTTAVWAFFAWTEVOITDHFHFQAITIXPFHXAFQHEFZQWGFLVWPTOFFA

Cryptanalysis II

- Can't try all 2^{88} simple substitution keys
- Can we be more clever?

- English letter frequency counts…



Cryptanalysis II

- Ciphertext:

PBFPVYFBQXZTYFPBFEQJHDXXQVAPTPQJKTOYQWIPBVWLXTOXBTFXQWAXBVCXQWAXFQJVWLEQNTOZQGGQLFXQWAKVWLX
QWAEBIPBFXFQVXGTVJVWLBTPQWAEBFPBFHCVLXBQUFEVWLXGDPEQVPQGVPPBFTIXPFHXZHVFAGFOTHFEFBQUFTDHZBQPO
THXTYFTODXQHFTDPTOGHFQPBQWAQJJTODXQHFOQPWTBDHHIXQVAPBFZQHCFWPFHPBFIPBQWKFABVYYDZBOTHPBQPQJTQ
OTOGHFQAPBFEQJHDXXQVAVXEBQPEFZBVFOJIWFFACFCCFHQWAUVWFLQHGFXVAFXQHFUFHILTTAVWAFFAWTEVOITDHFHFQA
ITIXPFHXAFQHEFZQWGFLVWPTOFFA

- Decrypt this message using info below

Ciphertext frequency counts:

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
21	26	6	10	12	51	10	25	10	9	3	10	0	1	15	28	42	0	0	27	4	24	22	28	6	8

Frequency analysis history

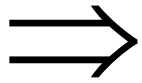
- Discovered by the Arabs
  - ✓ Earliest known description of frequency analysis is in a book by the 9-century scientist al-Kindi
- Rediscovered or introduced from the Arabs in Europe during the Renaissance
- Frequency analysis made substitution cipher insecure.

5. Double Transposition

- Plaintext: attackxatxdawn

	col 1	col 2	col 3
row 1	a	t	t
row 2	a	c	k
row 3	x	a	t
row 4	x	d	a
row 5	w	n	x

Permute rows and columns



	col 1	col 3	col 2
row 3	x	t	a
row 5	w	x	n
row 1	a	t	t
row 4	x	a	d
row 2	a	k	c

- Ciphertext: xtawxnattxadakc
- Key: matrix size and permutations (3,5,1,4,2) and (1,3,2)

One-Time Pad Encryption

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111

Encryption: Plaintext ⊕ Key = Ciphertext

P	h	e	i	l	h	i	t	l	e	r
	001	000	010	100	001	010	111	100	000	101
K	111	101	110	101	111	100	000	101	110	000
C	110	101	100	001	110	110	111	001	110	101
	s	r	l	h	s	s	t	h	s	r

One-Time Pad Decryption

Decryption:
Ciphertext ⊕ Key = (Plaintext ⊕ Key) ⊕ Key = Plaintext

C	s	r	l	h	s	s	t	h	s	r
	110	101	100	001	110	110	111	001	110	101
K	111	101	110	101	111	100	000	101	110	000
P	001	000	010	100	001	010	111	100	000	101
	h	e	i	l	h	i	t	l	e	r

One-Time Pad Summary

- **Provably secure**, when used correctly
  - ✓ Ciphertext provides no info about plaintext
  - ✓ All plaintexts are equally likely
  - ✓ Pad must be random, used only once
  - ✓ Pad is known only by sender and receiver
  - ✓ Pad is same size as message
  - ✓ No assurance of message integrity
- Why not distribute message(plaintext) the same way as the pad(key)?

One-Time Pad

- In terms of cryptography, the one-time pad is an encryption technique that cannot be cracked, but requires the use of a one-time pre-shared key the same size as, or longer than, the message being sent. In this technique, a plaintext is paired with a random secret key

One-Time Password

- A one-time password, also known as one-time pin, is a password that is valid for only one login session or transaction, on a computer system or other digital device.
- In short both are related to security but different content one was connect to authentication and another one is connected to encryption.

Stanford Univ. Cryptography 1 – Dan Boneh. 06. The One Time Pad

https://crypto.stanford.edu/~dabo/courses/OnlineCrypto/

Symmetric Ciphers: definition

Def: a **cipher** defined over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$

is a pair of "efficient" algs (***E***, ***D***) where

$E: \mathcal{K} \times \mathcal{M} \rightarrow \mathcal{C}$, $D: \mathcal{K} \times \mathcal{C} \rightarrow \mathcal{M}$

s.t. $\forall m \in \mathcal{M}, k \in \mathcal{K}: \; D(k, E(k, m)) = m$

- ***E*** is often randomized. ***D*** is always deterministic.

Consistency Equation : 모든 메시지, 모든 키에 대해 성립하고 암호화/복호화 과정은 정확해야 함

The One Time Pad (Vernam 1917)

First example of a "secure" cipher

$\mathcal{M} = \mathcal{C} = \{0,1\}^n$, $\mathcal{K} = \{0,1\}^n$

$\{0,1\}^n$ refers to the space of all n-length vectors consisting of 0s and 1s,

key = (random bit string as long the message)

메시지 M 길이만큼 동일한 길이의 키 K가 필요하고, 암호문 C도 동일한 길이를 가짐

Dan Boneh

참고 : $[0,1]^n$ (또는 $(0,1)^n$) refers to the space of all n-length vectors consisting of real numbers between 0 and 1 inclusive (또는 exclusive)

Stanford Univ. Cryptography 1 – Dan Boneh. 06. The One Time Pad

The One Time Pad (Vernam 1917)

$C := E(K,m) = K \oplus m$

$D(K,C) = K \oplus C$

msg:	0 1 1 0 1 1 1
key:	1 0 1 1 0 1 0 ⊕
CT:	

Indeed:

$D(K, E(K,m)) = D(K, K \oplus m) = K \oplus (K \oplus m) = (K \oplus K) \oplus m = 0 \oplus m = m$

Dan Boneh

You are given a message (m) and its OTP encryption (c).

Can you compute the OTP key from m and c?

No, I cannot compute the key.

Yes, the key is $k = m \oplus c$. ⇐

I can only compute half the bits of the key.

Yes, the key is $k = m \oplus m$.

Stanford Univ. Cryptography 1 – Dan Boneh. 06. The One Time Pad

The One Time Pad

Very fast enc/dec !!

… but long keys (as long as plaintext)

Is the OTP secure? What is a secure cipher?

What is a secure cipher?

Attacker's abilities: **CT only attack** (for now)

Possible security requirements:

attempt #1: **attacker cannot recover secret key**

$E(K, m) = m$ would be secure

attempt #2: **attacker cannot recover all of plaintext**

$E(K, m_0 \| m_1) = m_0 \| K \oplus m_1$ would be secure

Shannon's idea:

CT should reveal no "info" about PT

Dan Boneh

CT : Ciphertext, PT : Plaintext

Stanford Univ. Cryptography 1 – Dan Boneh. 06. The One Time Pad

Information Theoretic Security
(Shannon 1949)

Def: A cipher (E, D) over $(\mathcal{K}, \mathcal{M}, \mathcal{C})$ has **perfect secrecy** if

$\forall m_0, m_1 \in \mathcal{M}$ $(|m_0| = |m_1|)$ and $\forall c \in \mathcal{C}$

$$Pr[E(k, m_0) = c] = Pr[E(k, m_1) = c]$$

where k is uniform in $\mathcal{K}$ $(k \xleftarrow{R} \mathcal{K})$

⇒ given CT can't tell if msg is m_0 or m_1 (for all m_0, m_1)

⇒ most powerful adv. learns nothing about PT from CT

⇒ no CT only attack!! (but other attacks possible)

Information Theoretic Security

Def: A cipher ***(E,D)*** over (K,M,C) has **perfect secrecy** if

$\forall m_0, m_1 \in M$ $(|m_0| = |m_1|)$ and $\forall c \in C$

$$Pr[E(k, m_0) = c] = Pr[E(k, m_1) = c] \quad \text{where } k \xleftarrow{R} K$$

암호문 C로부터 평문 M을 찾아낼 확률이 모든 경우의 수에 대해 동일하게 일정함 (uniform)

암호문의 통계적 특성이 존재하지 않음 (M과 C 사이에 아무런 관계가 없음)

Stanford Univ. Cryptography 1 – Dan Boneh. 06. The One Time Pad

Lemma: OTP has perfect secrecy.

Proof:

$$\forall m, c: \quad \Pr_{k}[E(k,m)=c] = \frac{\#\text{keys } k \in \mathcal{K} \text{ s.t. } E(k,m)=c}{|\mathcal{K}|}$$

So: if $\forall m, c: \#\{k \in \mathcal{K} : E(k,m)=c\} = \text{const.}$

$\Rightarrow$ cipher has perfect secrecy

Let $m \in \mathcal{M}$ and $c \in \mathcal{C}$.

How many OTP keys map $\boldsymbol{m}$ to $\boldsymbol{c}$?

None

1 ⟸

2

Depends on $\boldsymbol{m}$

Stanford Univ. Cryptography 1 – Dan Boneh. 06. The One Time Pad

Lemma: OTP has perfect secrecy.

Proof:

For OTP: $\forall m, c$: if $E(k,m)=c$

$\Rightarrow k \oplus m = c \Rightarrow k = m \oplus c$

$\Rightarrow \boxed{\#\{k \in \mathcal{K} : E(k,m)=c\} = 1}$

$\Rightarrow$ OTP has perfect secrecy ∎

The bad news …

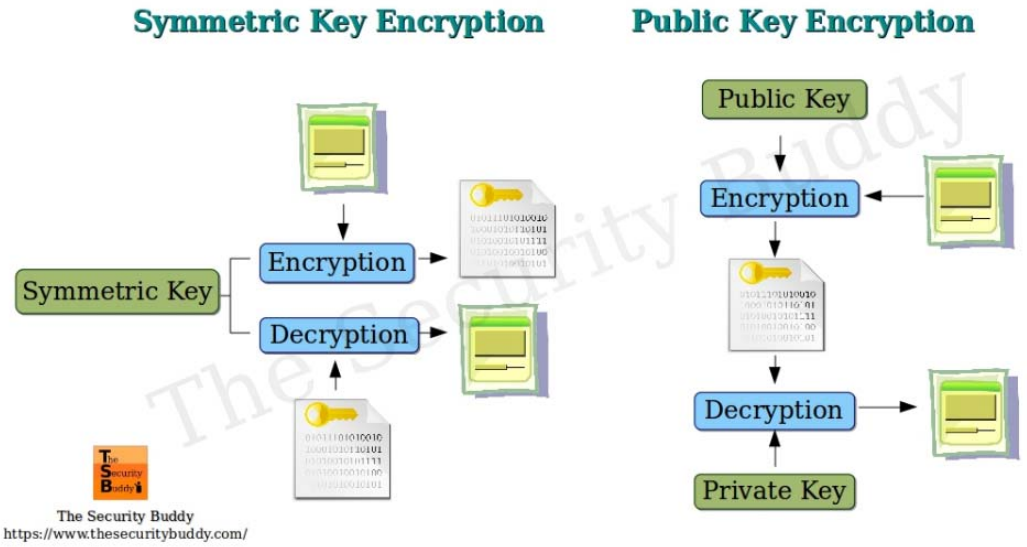
Thm: perfect secrecy $\Rightarrow$ $|\mathcal{K}| \geq |\mathcal{M}|$

i.e. perfect secrecy $\Rightarrow$ key-len $\geq$ msg-len

$\Rightarrow$ hard to use in practice !!

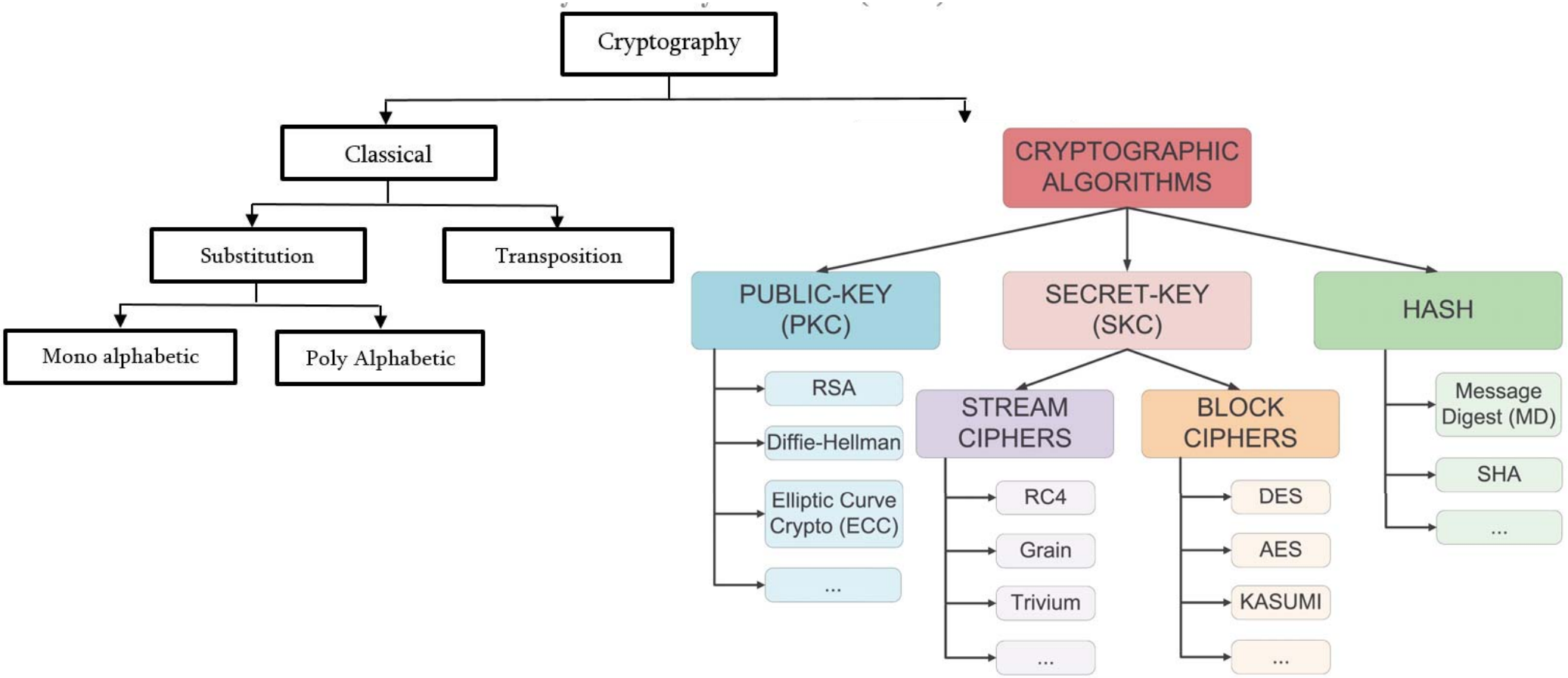
7. Taxonomy of Cryptography

- **Symmetric Key**
  - Same key for encryption as for decryption
  - Stream ciphers, Block ciphers
- **Public Key**
  - Two keys, one for encryption (public), and one for decryption (private)
  - Digital signatures — nothing comparable in symmetric key crypto
- **Hash algorithms**



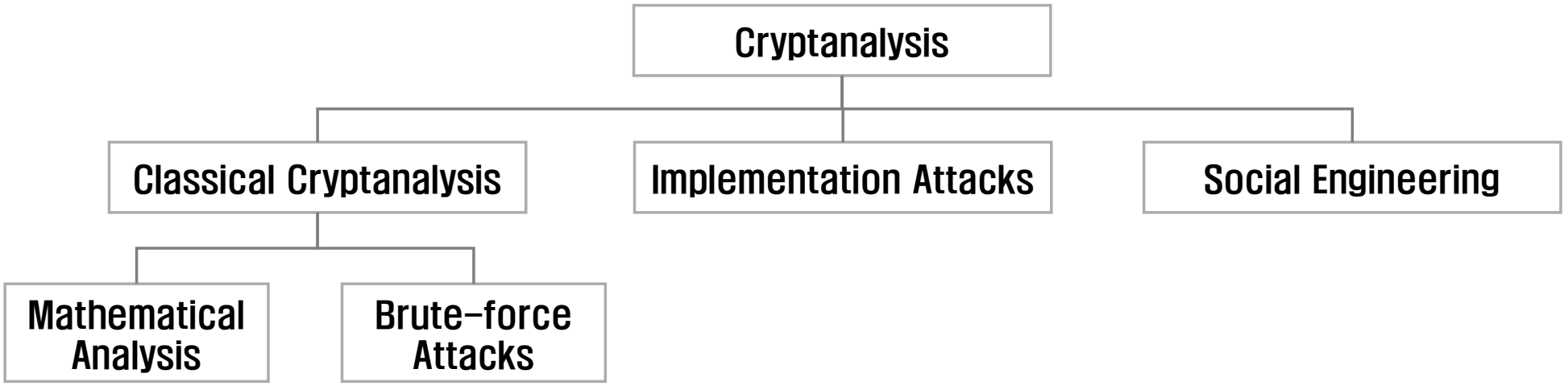
[그림 출처]
https://www.thesecuritybuddy.com/encryption/symmetric-key-encryption-vs-public-key-encryption/

7. Taxonomy of Cryptography



출처 : Why Quantum Cryptography Will Be The Future of Secure Communications?, https://medium.com/@bbsusheelkumar/a-deep-dive-into-quantum-cryptography-the-future-of-secret-communications-87f7bd2aad65

8. Taxonomy of Cryptanalysis



- Side-channel Analysis can obtain the key by measuring the electrical power consumption of a processor which operates on the key. The power trace can then be used to recover the key by applying signal processing techniques.
- Electromagnetic radiation can give information about the key.
- The implementation attacks are related to the physical access of attackers.

- Any human related activities such as bribing, blackmailing, tricking, or espionage
- Current fishing is a typical attack

8. Taxonomy of Cryptanalysis

- Ciphertext only (암호문 공격)
  - Algorithm and ciphertext only
- Known plaintext (알려진 평문 공격)
  - Some of plaintext and corresponding ciphertext
- Chosen plaintext (선택된 평문 공격)
  - Limited access to cryptosystem
  - "Lunchtime attack"
  - Protocols might encrypt chosen text
- Adaptively chosen plaintext (적응적 선택된 평문)
  - Choose the plaintext, View the resulting ciphertext,
    And choose the plaintext based the observed ciphertext
  - Related key
- Forward search (순방향 탐색)
  - (public key crypto only)
  - The case: plaintext is "yes" or "no"
- Etc., etc.

Safe key space

- How many keys are enough?

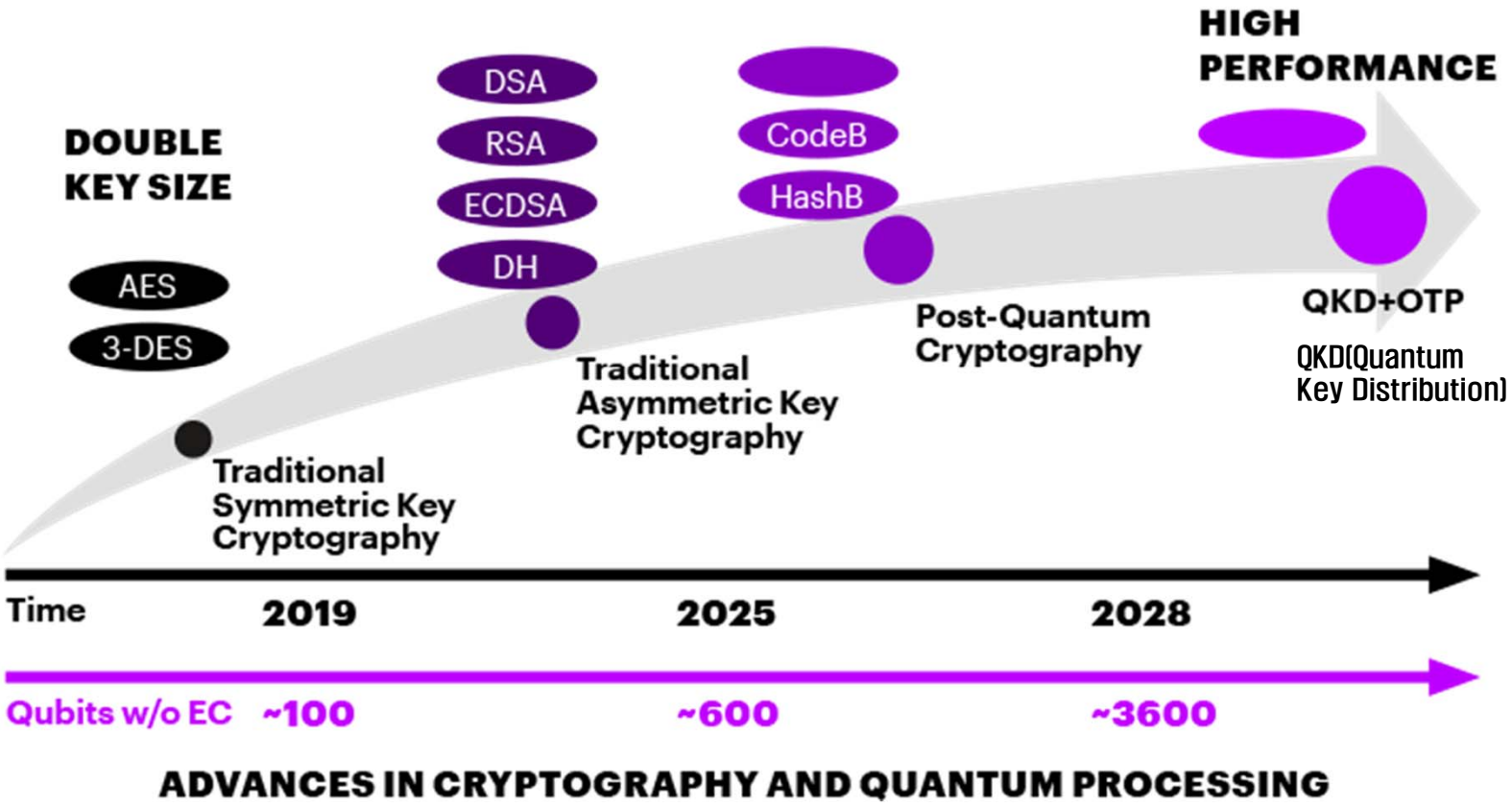
Key length	Estimated time
56–64 bits	A few hours or days
112–128 bits	Several decades in the absence of quantum computers
256 bits	Several decades even with quantum computers that run the current known quantum computing algorithms

Exhaustive key search times

Key size (bits)	Key space	Execution time (when 1 decrypt/us)	Execution time (when 10^6 decrypt/us)
32	$2^{32}=4.3\times10^{9}$	2^{31}us=35.8 min	2.15ms
56	$2^{56}=7.2\times10^{16}$	2^{56}us=1142yrs	10.01hrs
128	$2^{128}=4.3\times10^{38}$	2^{127}us=5.4×10^{24}yr	5.4×10^{18}yr
168	$2^{168}=4.3\times10^{50}$	2^{167}us=5.9×10^{36}yr	5.9×10^{30}yr

9. Advances in Cryptography

FIGURE 4. Timeline for future standardization events (copyright Accenture)



Existing cryptographic methods are the fabric of commerce, communications, identity and data protection at large—and all must be reviewed and potentially updated to continue conducting business safely and securely in a post-quantum world.

Every organization should work to achieve quantum-proof coverage by 2025.

출처 : Cryptography in a post-quantum world, https://www.accenture.com/il-en/insights/technology/quantum-cryptography