Security and Cryptography

2020-2021 Winter ABC Program

Team 5 Mentor 정현준

2021 / 02 / 08

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

문제!

100개의 약 중 하나가 독약이다

독약을 먹으면 1시간 후 죽는다

1시간이 주어졌을 때 독약을 찾기 위해 필요한

최소한의 토끼 수는?

(약을 먹이는데 걸리는 시간 무시, 약의 양 무한정)

정답!

CV #

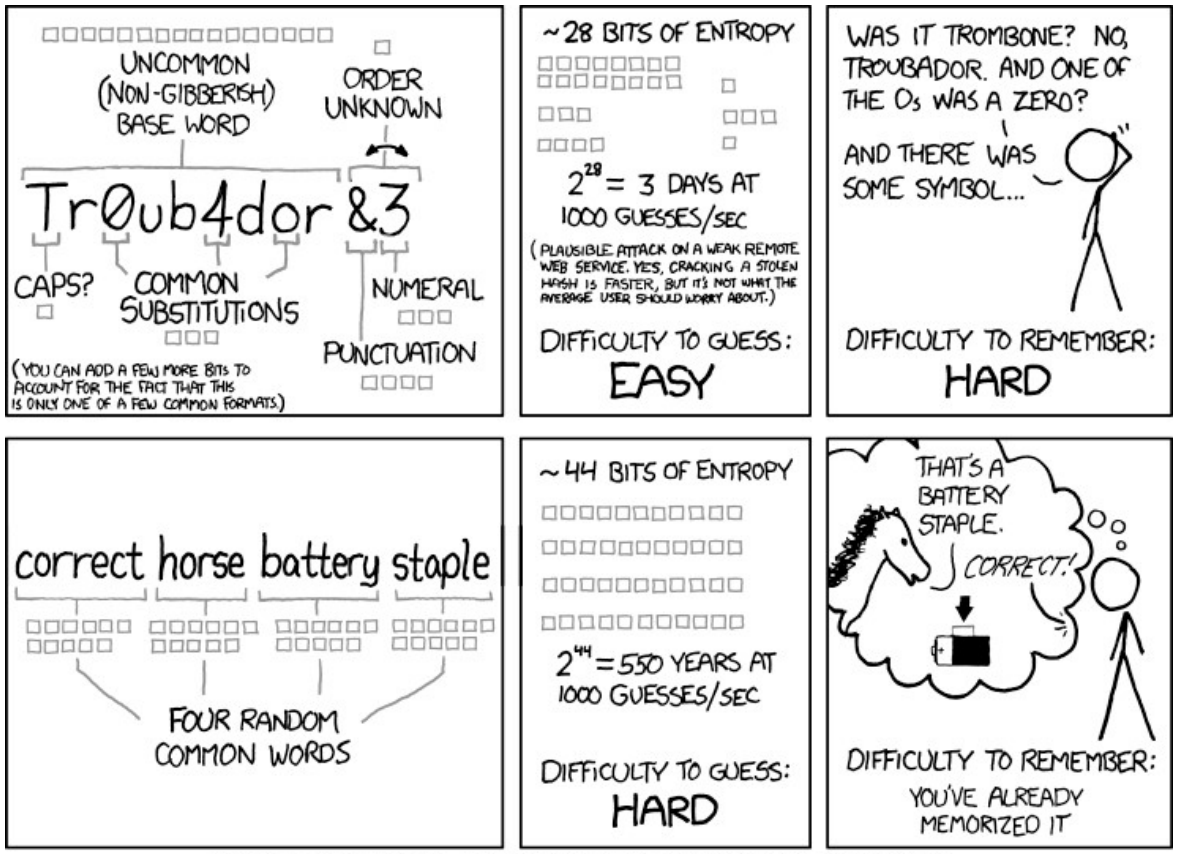
1번약 (이진수 0000001번약)은 1번 토끼 빼고 다 먹입니다.
2번약 (이진수 0000010번약)은 2번 토끼 빼고 다 먹입니다.
...
89번약 (이진수 1011001번약)은 1, 4, 5, 7번 토끼 빼고 다 먹입니다.
...
100번약 (이진수 1100100번약)은 3, 6, 7번 토끼 빼고 다 먹입니다.

만약 89번약이 독약이라면 1, 4, 5, 7번 토끼는 살고 나머지는 죽겠죠.
산 토끼를 1, 죽은 토끼를 0이라고 하면
7번토끼~1번토끼 순서대로 1011001 이 되고, 십진법으로 바꾸면 89가 나오게 됩니다.

Entropy

- 열역학적 엔트로피 -> 에너지의 무질서도
- 여기서는 정보 이론에서 말하는 엔트로피를 뜻합니다.
- 열역학에서의 엔트로피와 비슷하게 정보 이론에서의 엔트로피는 정보의 불확실성을 나타냅니다.
- 데이터의 단위로의 정보 엔트로피는 저장 또는 통신에 사용되는 평균 비트 수로 표현 됩니다.
- Ex) 매일 날씨의 정보가 2비트로 표현 될 수 있다면 하루의 날씨를 평균 2비트로 나타낼 수 있다.

Entropy


UNCOMMON (NON-GIBBERISH) BASE WORD
ORDER UNKNOWN
Tr0ub4dor&3
CAPS?
COMMON SUBSTITUTIONS
NUMERAL
PUNCTUATION
(YOU CAN ADD A FEW MORE BITS TO ACCOUNT FOR THE FACT THAT THIS IS ONLY ONE OF A FEW COMMON FORMATS.)
~28 BITS OF ENTROPY
2^28 = 3 DAYS AT 1000 GUESSES/SEC
(PLAUSIBLE ATTACK ON A WEAK REMOTE WEB SERVICE. YES, CRACKING A STOLEN HASH IS FASTER, BUT IT'S NOT WHAT THE AVERAGE USER SHOULD WORRY ABOUT.)
DIFFICULTY TO GUESS:
EASY
WAS IT TROMBONE? NO, TROUBADOR. AND ONE OF THE Os WAS A ZERO?
AND THERE WAS SOME SYMBOL...
DIFFICULTY TO REMEMBER:
HARD
correct horse battery staple
FOUR RANDOM COMMON WORDS
~44 BITS OF ENTROPY
2^44 = 550 YEARS AT 1000 GUESSES/SEC
DIFFICULTY TO GUESS:
HARD
THAT'S A BATTERY STAPLE.
CORRECT!
DIFFICULTY TO REMEMBER:
YOU'VE ALREADY MEMORIZED IT
THROUGH 20 YEARS OF EFFORT, WE'VE SUCCESSFULLY TRAINED EVERYONE TO USE PASSWORDS THAT ARE HARD FOR HUMANS TO REMEMBER, BUT EASY FOR COMPUTERS TO GUESS.

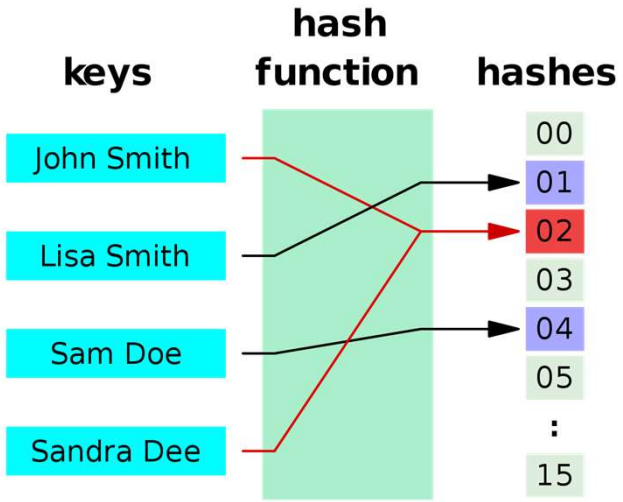
용어

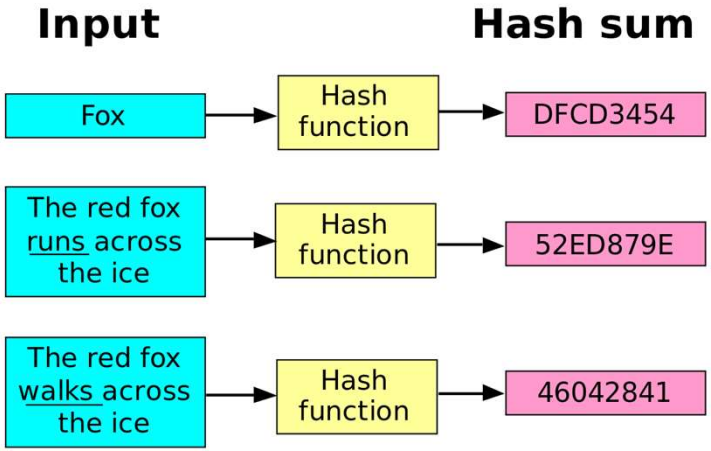
- 키 (key)
- 평문 (plaintext, message)
- 암호문 (ciphertext)
- 암호화 (encryption)
- 복호화 (decryption)

시저 암호 예시

Hash Function

- 임의의 길이를 가진 데이터를 고정된 크기의 데이터로 특정한 규칙을 통해 매핑하는 역할을 합니다.
- 이는 암호화랑은 차이가 있습니다. (복호화가 없기 때문입니다.)
- 대표적으로 Git에서 사용하였던 SHA-1 Hash가 있습니다.



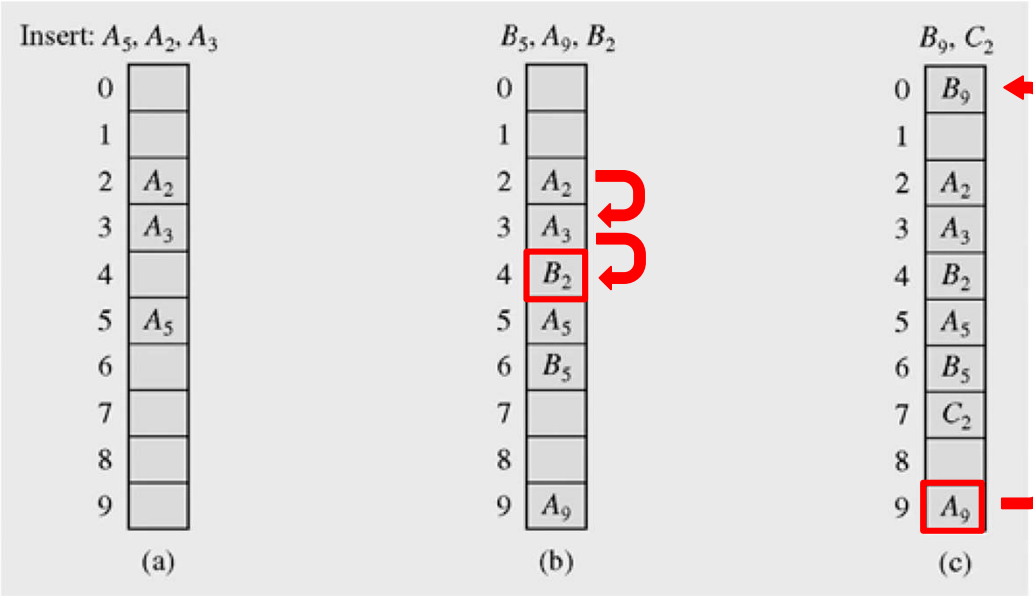


Hash Function

- Deterministic : 같은 input을 hash function에 넣으면 결과는 항상 같습니다.
- Non-invertible : output을 input으로 되돌리는 방법이 없습니다. (단방향)
- Hash collision : 여러 개의 input이 동일한 값으로 매핑 되는 경우를 말합니다.
- Ex) hash_function(input) = input % 10
  - Input은 10의 나머지 값을 hash function의 output으로 매핑 됩니다.
  - 이 경우, 11과 121 input은 같은 1 값으로 매핑이 됩니다.
  - Hash collision은 이러한 상황을 의미합니다.
- Hash collision을 해결하는 방법으로는 크게 open-addressing과 chaining이라는 방법이 일반적으로 많이 사용됩니다.

Open addressing

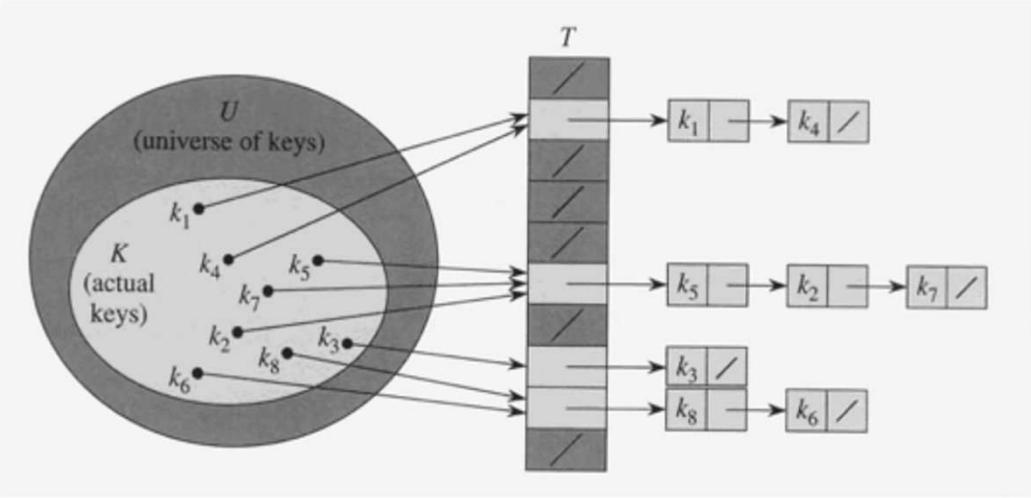
- 충돌이 일어나는 값이 들어온 경우, 정해진 policy에 따라 다른 값에 매핑을 하게 됩니다.
  - Linear probing, Quadratic probing 등등…



Linear probing 예시

Chaining

- 동일 장소로 해싱된 모든 값들을 하나의 linked list로 저장합니다.
- Open addressing에서 발생할 수 있는 충돌이 발생하지 않습니다.
- 모든 값들이 하나의 output으로 해싱되는 경우 검색에 많은 시간이 필요할 수 있습니다.

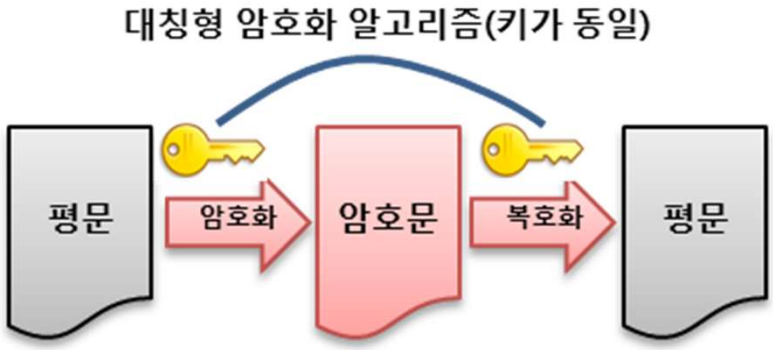


Key Derivation Function (KDF)

- 키 유도 함수라고 부릅니다.
- Master key나 어떠한 비밀 정보로부터 특정 시스템이나 네트워크의 암호화, 또는 복호화에 필요한 암호 키를 유도하는 함수를 의미합니다.
- 즉, 키 값을 만드는 일종의 키 생성기(key-gen), 혹은 난수 생성기 입니다.
- Ex) ssh-keygen 명령어를 사용하면 passphrase를 묻습니다.
  - ssh-keygen은 입력한 passphrase를 기반으로 KDF를 실행합니다.
  - KDF로부터 private key와 public key가 생성됩니다.

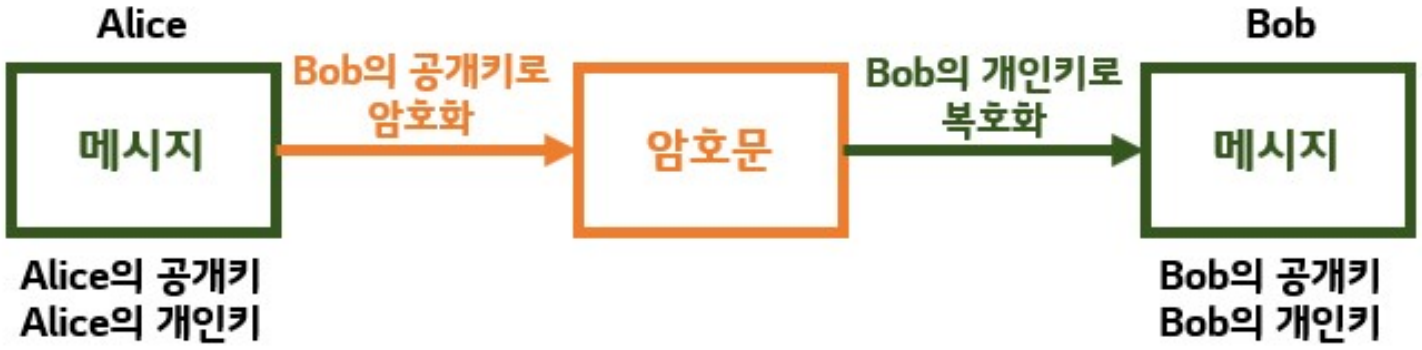
Symmetric Cryptography

- 암호화와 복호화에 동일한 키가 사용 됩니다.
- 비대칭 알고리즘에 비해 암호화/복호화 속도가 빠릅니다.
- 주로 데이터 통신에서 사용 됩니다. (통신에서는 속도가 중요하기 때문)
- 동일한 키를 암호화와 복호화에 사용하기 때문에 공격자가 키를 가로채면 암호가 쉽게 풀린다는 단점이 있습니다.



Asymmetric Cryptography

- 비대칭 암호화 방식이라고 합니다. (또는 공개키 암호화 방식)
- Public key와 Private key를 이용해 암호화와 복호화를 수행합니다.
- Public key => 암호화에 사용, Private key => 복호화에 사용
- 보안 강도가 높은 만큼 연산이 복잡하고 느립니다.



Assignment

- 아래 링크에서 동영상을 보고 옵시다.
- https://missing.csail.mit.edu/2020/potpourri/
- 이전 카테고리에서 다루지 못한 여러가지를 다룹니다.
- 내용이 조금 많기 때문에 다음 시간에는 API, Markdown, Docker, VM, Cloud 위주로 할 생각입니다.
- (위에 해당하는 내용만 보고 오셔도 됩니다!)

Thank you


UNIST
ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY
2009