2. 생일 문제와 해시함수

해시함수 충돌 저항성은 생일 문제와 밀접한 연관을 가진다.
생일 문제는 확률론에서 크게 4가지가 있다.
여기서는 해시함수와 연관된 2가지를 소개한다.

1. 첫 번째 생일 문제 - 생일 패러독스(paradox, 역설) : 강한충돌저항성과 연관

문제	생일이 같은 사람이 적어도 한 쌍 이상 존재할 확률이 50% 이상이 되기 위해서는 몇 사람이 필요한가?

- 먼저, 정답을 말하면 **23명**이다. - 생각보다 **적은** 인원이다.
- 평균적으로, 23명을 조사하면 생일이 같은 한 쌍이 존재할 확률이 50% 이상이다.
- 생각보다 적은 인원인 23명을 조사하면 생일이 같은 한 쌍이 존재할 확률이 50% 이상이다.
- 너무 놀라워서, **생일 패러독스(역설) 문제**라고 한다.
- 해시함수의 강한충돌저항성은 같은 출력을 내는 임의의 서로 다른 두 입력을 찾아내는 계산을 논하는 것이다.
- 해시함수의 **강한충돌저항성**은 생일 패러독스 문제와 연관된다.(뒤에서 구체적으로 다룬다)

2. 두 번째 생일 문제 - 약한충돌저항성과 연관

문제	교실에 k명의 학생이 있을 때, 적어도 한 학생이 교사가 미리 지정한 학생과 생일이 같을 확률이 50% 이상이 되기 위한 k의 최소값은?

- 먼저, 정답을 말하면 **254명**이다.
- 어떤 사람이 나와 생일이 같을 확률은 1/365이다. → 1년을 365일로 간주
- 확률이 1이 되기 위해서는 교실 안에 365명이 있어야 한다.(미리 지정한 학생 포함)
- 확률이 50% 이상 되기 위해서는(확률 P, 인원 k, N = 365)
- 답을 구하면, **254명**이다.
- 두 번째 생일 문제는 해시함수의 **약한충돌저항성**과 연관된다.

기출문제 분석

1. 생일 역설(Birthday Paradox)에 대한 설명으로 옳지 않은 것은? [2014년 국가 7급]

① 해시함수(hash function)는 충돌 메시지 쌍을 찾아내는 데 사용된다.
② 특정 장소에서 23명 이상이 있으면, 그 중에서 2명 이상의 사람이 생일이 같을 확률은 0.5보다 크다.
③ 블록 암호 알고리즘의 안전성을 분석하는 데 이용된다.
④ 0부터 N－1까지의 균일분포를 갖는 수 중에서 임의로 한 개의 수를 선택한다면, $(N)^{1/2}$ 번의 시도 후에 동일한 수가 반복해서 선택될 확률은 0.5를 넘는다는 이론과 부합한다.

♤ **생일 역설 – 해시함수의 강한 충돌 저항성**

- 생일 역설은 암호학적 **해시함수** 알고리즘의 안전성을 분석하는 데 이용된다.

// 생일 역설 : 생일이 같은 사람이 적어도 한 쌍 이상 존재할 확률이 50% 이상되려면, 23명이 필요하다.
↓증명

- 임의 정수가 1~n까지 균등확률로 분포되어 있고, 이중에서 임의로 k개(k≤n)를 선택하였을 때,
- 적어도 한 쌍이 중복되어 나타날 확률을 P(n, k)라고 하면,
- $P(n, k) = 1 - \frac{n!}{(n-k)! \times n^k}$ 로 표현된다.
- 여기서, $P(n, k) \geq \frac{1}{2}$ 인 k를 계산하면, → 50% 확률

  $k = \sqrt{2(\ln 2)n} = 1.17\sqrt{n} \simeq \sqrt{n}$ 이다.
- 생일 패러독스에서 n=365이므로 $k = 1.17\sqrt{n} = 1.17 \times \sqrt{365} \simeq$ **23(명)**

정답 : ③

2. 해시함수의 충돌저항성을 위협하는 공격 방법은? [2020년 지방 9급]

① 생일공격 ② 사전공격
③ 레인보우 테이블 공격 ④ 선택평문공격

♤ **해시함수의 충돌저항성**

- 생일 패러독스(paradox, 역설) : 강한충돌저항성과 연관
- 평균적으로, 23명을 조사하면 생일이 같은 한 쌍이 존재할 확률이 50% 이상이다.

정답 : ①