NTIN066 - solutions 8

David Mareček

April 06, 2022

1. Prove that if H is (k,c)-independent for $k > 1$, then it is also $(k-1,c)$-independent

$$P[h(x_1) = a_1 \wedge h(x_2) = a_2 \wedge \cdots \wedge h(x_{k-1}) = a_{k-1}] =$$

$$= \sum_{i=1}^{m} P[h(x_1) = a_1 \wedge h(x_2) = a_2 \wedge \cdots \wedge h(x_{k-1}) = a_{k-1} \wedge h(x_k) = i] \leq$$

$$\leq m\frac{c}{m^k} = \frac{c}{m^{k-1}}$$

2. Prove that if H is $(2,c)$-independent, then it is c-universal.

$$\forall x_1, x_2, a_1, a_2 : P[h(x_1) = a_1 \wedge h(x_2) = a_2] \leq \frac{c}{m^2}$$

$$P[h(x_1) = h(x_2)] = \sum_{i=1}^{m} P[h(x_1) = i \wedge h(x_2) = i] \leq m\frac{c}{m^2} = \frac{c}{m}$$

3. Prove tha tabulation hashing is 3-independent but not 4-independent.

If tabulation uses only one table, it is perfectly random, so it is k-independent for each k not exceeding the size of the universe.

Otherwise, we have $t \geq 2$ tables, each indexed by bits. Now consider x_1, x_2, x_3 and a_1, a_2, a_3 from the definition of 3-independence. At the same time, we imagine x_i as ordered k-tuples of b-bit values (parts indexing individual tables).

(a) If there is a position i in which all x_1, x_2, x_3 differ, then regardless of the values in the other positions, we can always fill in the table corresponding to this position so that the XORs come out a_1, a_2, a_3. Since T_i is perfectly random, the probability of generating a_1, a_2, a_3 is $1/m^3$.

(b) Or there exists (WLOG) a position i and position j, for which:

```
x1[i]  x2[i],  x1[i] = x3[i],    (call x1[i] = A and x2[i] = B)
x1[j] = x2[j],  x1[j]  x3[j],    (call x1[j] = C and x3[j] = D)
```

```
We get a set of equations (^ is XOR)

Ti[x1[i]] ^ Tj[x1[j]] ^ v1 = a1
Ti[x2[i]] ^ Tj[x2[j]] ^ v2 = a2
Ti[x3[i]] ^ Tj[x3[j]] ^ v3 = a3

where v1, v2, and v3 are XORed values from the other tables.

Ti[A] ^ Tj[C] = a1 ^ v1
Ti[B] ^ Tj[C] = a2 ^ v2
Ti[A] ^ Tj[D] = a3 ^ v3
```

For each $T_j[C]$ and each v_1, v_2, v_3, this set of equations has just one solution. So the probability that filling of the tables leads to a solution is exactly $1/m^3$, as we need for 3-independence.

```
Counter-example for 4-independence:
Imagine these four hash keys.
Values A, B, C, D are corresponding values in the hash tables.

A B E F
A C E F
D B E F
D C E F

Then for each a1, a2, a3, a4:

A ^ B ^ E ^ F = a1
A ^ C ^ E ^ F = a2
D ^ B ^ E ^ F = a3
D ^ C ^ E ^ F = a4

=> a1 ^ a2 ^ a3 ^ a4 = 0

If a1 ^ a2 ^ a3 ^ a4 = 0, probability is 1/m3
Otherwise, probability is 0.
```