

# NSUCRYPTO2024

## Problem 11: A simple hash function

October 21, 2024

### Solution

#### Collision algorithm

To ensure that the input sequence  $P$  is compatible with the hash function, we employ the following padding strategy:

1. **Case 1:  $|P|$  is a multiple of 6:**

- We append six zeros to the sequence  $P$  to obtain a new padded sequence  $P' = P \parallel 000000$ , where  $\parallel$  denotes concatenation.

2. **Case 2:  $|P|$  is not a multiple of 6:**

- We pad the sequence by appending the first digit of the predefined padding sequence "12345". For example, if  $P$  is 7256, it should be modified to  $P' = 725612$ .

#### Proof of Correctness

We need to prove that the padding algorithm correctly preserves the hash function  $H$  for both cases.

**Case 1: Length of  $P$  is a multiple of 6**

In this case, we pad  $P$  with six zeros:

$$H(P', K) = H(P \parallel 000000, K)$$

Using the properties of the hash function, we can express this as:

$$H(P', K) = H(P, K) + (-1)^{\frac{|P|}{6} \pmod{2}} \cdot H(000000, K) = H(P, K) + 0 = H(P, K)$$

This shows that the hash of the modified sequence  $H(P', K)$  is equal to the hash of the original sequence  $H(P, K)$  when  $|P|$  is a multiple of 6, independent of the key  $K$ .

## Case 2: Length of $P$ is not a multiple of 6

In this scenario, we pad  $P$  with the digits from the predefined padding sequence:

$$P' = \text{pad}(P)$$

Thus, we have:

$$H(P', K) = H(P, K)$$

Since the padded modified sequence equals the padded original sequence, the hash value remains unchanged.

## Conclusion

In both cases, we have shown that:

$$H(P', K) = H(P, K)$$

This proves the correctness of the collision algorithm for the hash function.

## Shortest Collision for $P = 134875512293$

We claim that the shortest collision for  $P$ , denoted as  $\min(|P'|)$ , such that  $H(P', K) = H(P, K)$ , is 7. The modified sequence  $P'$  can take values such as:

$$P' \in \{1349275, 2349276, 3349277, 4349278, 5349279\}$$

To demonstrate that  $|P'| = 7$  is indeed the shortest length for which a collision occurs, we will systematically explore all possible padded sequences  $P'$  of varying lengths  $l$  such that  $H(P', K) = H(P, K)$ , where  $l = 1, 2, \dots$

The Python code utilizes the SageMath library to implement this approach. For more details, please refer to our GitHub repository.

## Explanation of the Code

1. **Padding Function:** The `pad` function takes an input sequence  $P$  and appends necessary digits from the predefined padding sequence until the length of  $P$  becomes a multiple of 6.
2. **Hash Function:** The `hash` function computes the hash value  $H$  of the padded sequence  $P$  based on the key  $K$ . It divides  $P$  into blocks of 6 digits and computes the hash using a specific formula involving the key.

3. **Collision Search:** The loop iterates through lengths  $l$  from 1 to 7, calculating the hash values for all possible keys  $K$  (from 0 to 63). It then checks if a collision occurs by comparing the computed hash values.
4. **Matrix Operations:** It sets up a matrix equation to find the solution vector  $v_0$ . If the matrix is singular, indicating no solutions exist for that length, it catches the `ValueError` and continues.

This method allows us to systematically identify whether there are any padded sequences  $P'$  of length  $l < 7$  that produce the same hash as  $H(P, K)$ . If none are found for lengths 1 through 6, we can conclude that  $|P'| = 7$  is indeed the shortest length for which a collision exists.

To analyze the results obtained from the code execution for lengths  $l = 6$  and  $l = 7$ :

For  $l = 6$ , the output is:

$$[-4, 2, 2, 6, -2, 2]$$

However, this solution does not meet the requirement since the output values must be digits from 0 to 9. The presence of negative values  $-4$  and  $-2$  indicates that this solution is invalid.

In contrast, for  $l = 7$ , the output is:

$$[-4, 3, 4, 9, 2, 7, 0]$$

Here, both  $-4$  and  $0$  are the first elements of even/odd blocks, with the first block being even and the second being odd. This property allows us to increment each value by a constant  $k$  without changing the overall hash value  $H$ .

From this, we can derive valid padded sequences  $P'$  by adjusting the values appropriately. For example, if we set  $k = 4$ , the adjusted values yield:

$$0349274, 1349275, 2349276, 3349277, 4349278, 5349279$$

Thus, we conclude that valid padded sequences of length  $|P'| = 7$  can be generated by incrementing the elements accordingly, while ensuring the hash value remains unchanged.

Please refer to the solution script for more details on [NSUCRYPTO2024 Problem 11](#).