Assignment 6 Security

Mohamed Tarek Aiad	19016405
Michael Samir Azmy	19016237
Mohamed Momen Salama	19016474

Code Explanation:

- **simple_hash:**
  This function calculates a hash value for a given file. It reads the file byte by byte, circularly shifts the hash value to the left by four positions, and XORs the new byte with the least significant byte of the hash value. Finally, it returns the resulting hash value.

- **hash_all_files:**
  This function calculates hash values for all files within a given directory. It iterates over each file in the directory, calculates its hash value using the '**simple_hash**' function, and stores the results in a dictionary where the keys are file names and the values are their corresponding hash values.

- **find_duplicate_hashes:**
  This function finds duplicate files based on their hash values. It reads the hash values from an input file, maintains a dictionary where hash values are keys and file paths are values, and identifies duplicate hash values. If a hash value is already present in the dictionary, it indicates a duplicate file. The function returns a list of tuples, where each tuple contains the original file path and the path of its duplicate.

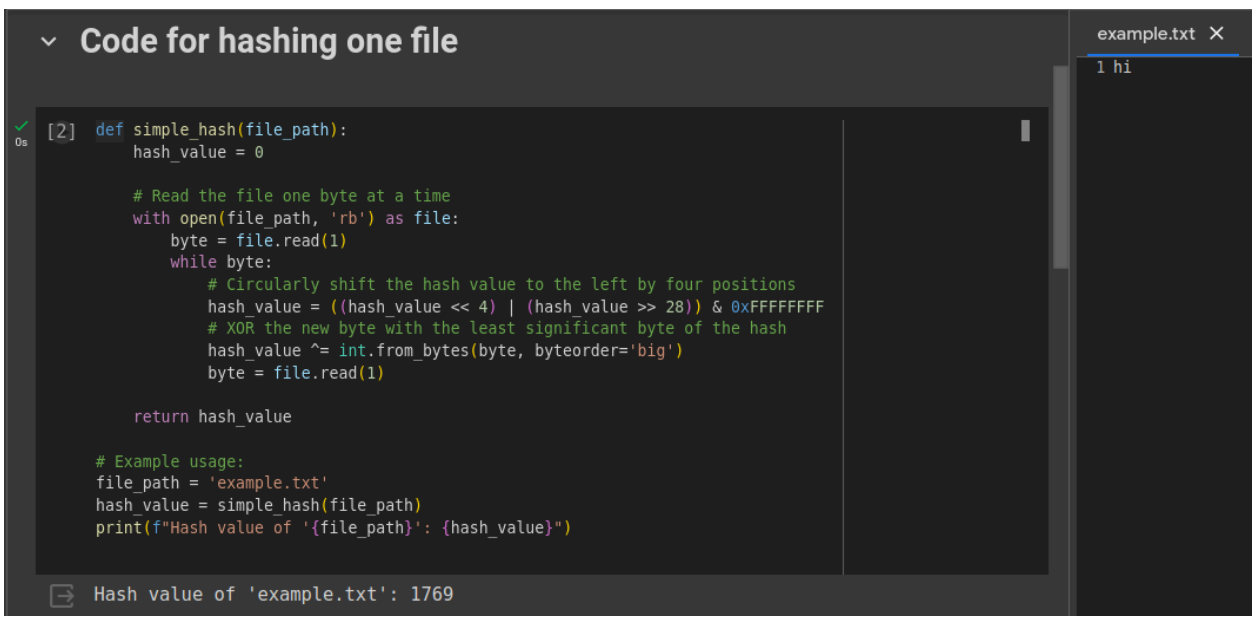
- **Make_collision:**
  This function aims to generate collisions between two files. The function takes file paths as input, reads the content of the second file, and calculates hash values for both files. By performing a bitwise XOR operation on the hash values, it generates an XOR hash, which is then converted into bytes and extended to match the length of the second file's content. The function combines the file content with the extended XOR bytes to create a collision file. Afterward, it calculates the hash of the collision file and prints the original hash values for the input files alongside the collision file hash, indicating whether a hash collision has been detected.

How to use :

- To hash a file, just call the method ‘**simple_hash**’ with the file.
- To hash set of files, call the function ‘**hash_all_files**’ and give it the directory that contains the files.

Sample runs:

- Try to hash a file named ‘example.txt’ with text ‘hi’ in it.



- Try to hash all the assignment files, so i made a directory and put the files in. The results are saved in ‘hash_values.txt’.

Hash group of files

```python
[5] import os

def hash_all_files(directory):
    hashes = {}
    for file_name in os.listdir(directory):
        file_path = os.path.join(directory, file_name)
        if os.path.isfile(file_path):
            hash_value = simple_hash(file_path)
            hashes[file_name] = hash_value
    return hashes

# Example usage:
directory = '/content/drive/MyDrive/Ass6_sec'
output_file = 'hash_values.txt'
hashes = hash_all_files(directory)
with open(output_file, 'w') as f:
    for file_name, hash_value in hashes.items():
        f.write(f"{file_name}: {hash_value}\n")
print("Hash values dumped into", output_file)
```

```
Hash values dumped into hash_values.txt
```

hash_values.txt

```
Assignment 1.pdf: 2670865059
Assignment 2.pdf: 2781405567
Assignment 4.pdf: 275180873
Assignment 5.pdf: 1164850897
Assignment 6.pdf: 1621738671
```

- Check for duplicates in the previous hash files.

Check duplicates

```python
[7] def find_duplicate_hashes(output_file):
        hash_dict = {}
        duplicates = []
        with open(output_file, 'r') as file:
            for line in file:
                file_path, hash_value = line.strip().split(':')
                if hash_value in hash_dict:
                    duplicates.append((hash_dict[hash_value], file_path))
                else:
                    hash_dict[hash_value] = file_path
        return duplicates

output_file = 'hash_values.txt'
duplicates = find_duplicate_hashes(output_file)

if duplicates:
    print("Duplicate files found:")
    for duplicate in duplicates:
        print("Original file:", duplicate[0])
        print("Duplicate file:", duplicate[1])
else:
    print("No duplicate files found.")
```

```
No duplicate files found.
```

hash_values.txt

```
Assignment 1.pdf: 2670865059
Assignment 2.pdf: 2781405567
Assignment 4.pdf: 275180873
Assignment 5.pdf: 1164850897
Assignment 6.pdf: 1621738671
```

- ## Make collision

```python
# get file to make a collision
def make_collision(file1, file2):

    with open(file2, "rb") as file:
        file_content = file.read()

    hash_file1_original = simple_hash(file1)
    hash_file2_original = simple_hash(file2)

    xor_hash = hash_file1_original ^ hash_file2_original

    xor_bytes = xor_hash.to_bytes(4, byteorder='big')

    extended_xor_bytes = b''
    for i in range(4):
        extended_xor_bytes += bytes([0]) + xor_bytes[i:i+1]

    collision_file_content = file_content + extended_xor_bytes

    with open("collision_file.txt", "wb") as collision_file:
        collision_file.write(collision_file_content)

    collision_file_hash = simple_hash("collision_file.txt")

    print("Original Hash for file1.txt:", hash_file1_original)
    print("Original Hash for file2.txt:", hash_file2_original)
    print("Collision file hash         ", collision_file_hash)
    print("Hash collision detected:    ", hash_file1_original == collision_file_hash)
    print()

make_collision("file1.txt", "file2.txt")
```

```
Original Hash for file1.txt: 573866766
Original Hash for file2.txt: 803974021
Collision file hash          573866766
Hash collision detected:     True
```

File1 content:

Walt, I’ve said it before, if you are in danger, we go to the police. I do not say that lightly. I know what it could do to this family. But if it’s the only real choice we have, if it’s either that or you getting shot when you open your front door. You’re not some hardened criminal, Walt, you are in over your head. That’s what we tell them. And that’s the truth. A school teacher, cancer, desperate for money, roped into working, unable to even quit. You told me that yourself, Walt. Jesus, what was I thinking? Walt, please! Let’s both of us stop trying to justify this whole thing and admit you’re in danger.

File2 content:

Who are you talking to right now? Who is it you think you see? Do you know how much I make a year? I mean, even if I told you, you wouldn't believe it. Do you know what would happen if I suddenly decided to stop going into work? A business big enough that it could be listed on the NASDAQ goes belly up, disappears. It ceases to exist without me. No, you clearly don't know who you're talking to so let me clue you in. I am not in danger Skylar, I am the danger. A guy opens his door and gets shot and you think that of me? No, I am the one who knocks

Code:

Assignment6_Security.ipynb