#### **NILADRI MITRA**

**VIT REG.NO: 20BIT0381** 

**COURSE:** CYBERSECURITY & ETHICAL HACKING(SMARTBRIDGE EXTERNSHIP)

Vit email id: niladri.mitra2020@vitstudent.ac.in

# <u>Digital Assignment – 3</u>

# <u>Symmetric Key Algorithm – AES(Advanced Encryption Standard)</u>

Advanced Encryption Standard (AES) is a popular symmetric encryption algorithm that protects sensitive data. The National Institute of Standards and Technology (NIST) chose it in 2001 to replace the outdated Data Encryption Standard (DES). AES offers high security and efficacy, making it appropriate for various applications. AES operates on data blocks of fixed length, with a block size of 128 bits. It employs a key measure of 128, 192, or 256 bits, depending on the required level of security. The algorithm consists of multiple transformation cycles involving key mixing, substitution, and permutation.

The critical expansion procedure generates round keys from the original encryption key. Each successive round key derives from its predecessor and undergoes a sequence of transformations. This assures that each round key is distinct and provides substantial randomness. AES performs a series of cycles on the input data during encryption. Each round includes four fundamental operations: SubBytes, ShiftRows, MixColumns, and AddRoundKey.

SubBytes replaces each byte in the data with a value from a substitution box (S-box). This non-linear substitution adds confusion and increases the efficacy of the encryption.

The ShiftRows operation adjusts the bytes in each row of the data matrix cyclically, providing diffusion and spreading the data across the entire block.

The MixColumns operation applies a mathematical transformation to each column of the data matrix to combine the bytes. This operation adds diffusion and further strengthens the security of the encryption.

The AddRoundKey operation combines the current round's round key with the data. It combines the bits using a bitwise XOR operation, effectively incorporating the key material into the encryption process.

The ultimate encrypted data is obtained after a predetermined number of rounds (depending on the key size). AES decryption follows a similar procedure, albeit with inverse operations for SubBytes, ShiftRows, and MixColumns, and with the round keys applied in reverse sequence.

Due to its robust security properties and efficient implementation on modern hardware, AES has seen widespread adoption. It has become the de facto standard for symmetric encryption. It is used in various applications, including securing communications, safeguarding sensitive data, and guaranteeing the integrity of digital content.

#### **Advantages**

- i. **Robust Security:** AES is highly secure and withstood extensive cryptanalysis. The cryptographic community has rigorously tested and evaluated it, making it a reliable option for protecting sensitive data.
- ii. **Efficiency:** AES is efficient concerning both computational and memory resources. It is optimised for contemporary processors, enabling quick encryption and decryption operations.
- iii. **Versatility:** AES supports key sizes of 128, 192, and 256 bits, providing the flexibility to select the desired level of security. It can accommodate a variety of applications, ranging from the protection of small quantities of data to the protection of large systems.
- iv. **Resistance to Attacks:** AES is designed to withstand various attacks, such as differential and linear cryptanalysis, brute-force attacks, and known-plaintext attacks. In addition to the key expansion, its substitution and permutation operations contribute to its resistance to cryptographic attacks.

#### **Disadvantages**

- i. **Side-Channel Attacks:** AES implementations are vulnerable to side-channel attacks, which exploit information revealed during the encryption process, such as timing or power usage. These assaults may enable the recovery of the secret key. To ensure secure use, implementations must meticulously mitigate side-channel vulnerabilities.
- ii. **Key Management:** AES's security is significantly dependent on the encryption key. Weak key management practices, such as using brief or predictable keys, can compromise AES's security. It is essential to use strong and correctly managed keys to prevent attacks.
- iii. **Related-Key Attacks:** AES is susceptible to related-key attacks, in which an adversary exploits the relationship between various keys to recover information or weaken the encryption. However, related-key assaults are deemed impractical in the majority of real-world circumstances.

- iv. **Quantum Computing Threat:** As with most classical cryptographic algorithms, AES is potentially vulnerable to quantum computer-based attacks. Quantum computers may be able to crack the symmetric encryption used by AES by performing a brute-force search of the key space efficiently. However, it is essential to note that practical, large-scale quantum computers capable of breaching AES have yet to be made available.
- v. **Implementation Flaws:** Vulnerabilities may result from implementation flaws or deficiencies rather than the algorithm itself. AES's security can be compromised by errors in implementation, incorrect utilisation, or insufficient configurations.

### **Real World Examples**

- i. **VPNs (Virtual Private Networks):** VPNs use AES encryption to securely connect you to another server online so that your data does not escape.
- ii. Numerous prominent apps (such as Snapchat and Facebook Messenger) use AES encryption to transmit data such as photos and messages securely.
- iii. **Archive and compression tools:** All major file compression applications employ AES to prevent data leakage. These applications consist of 7z, WinZip, and RAR.
- iv. **OS system components:** AES is utilised to protect the confidentiality of data on wireless networks, such as Wi-Fi networks, using encryption. Additionally, it can be used to encrypt sensitive data stored in databases. This protects sensitive information.

### <u>Asymmetric Key Algorithm – RSA</u>

The RSA algorithm, devised in 1977 by Ron Rivest, Adi Shamir, and Leonard Adleman, is a popular asymmetric encryption algorithm. RSA enables secure communication by allowing the encryption and decryption of data using a pair of public and private keys. Large composite numbers are challenging to factor into their prime factors, contributing to the algorithm's security.

#### RSA functions as follows:

- i. The initial process involves generating the RSA key pair. It requires choosing p and q, two enormous prime numbers. These figures are kept confidential. n, the product of p and q, is the modulus for the keys. As the public exponent, another value, e, is typically a small prime number, such as 65537. d is computed using modular arithmetic such that (d \* e) mod ((p 1) \* (q 1)) = 1.
- ii. To encrypt a message, the sender must obtain the recipient's public key, composed of the modulus n and the public exponent e. The sender converts the plaintext message to a number less than n. The process of encryption entails computing c = (me) mod n. The resultant ciphertext, c, is transmitted to the intended recipient.
- iii. Decryption is carried out by the recipient, who possesses the corresponding private key with exponent d. The recipient calculates m as (cd) divided by n. This operation recovers the message's original plaintext.
- iv. RSA can be used for digital signatures as well. To sign a message, the sender uses their private key to calculate its hash value and then encrypts the hash with the private exponent d. The recipient can authenticate the signature by decrypting it with the sender's public key and comparing the resulting hash value to a newly calculated hash of the received message.

RSA is secure due to the difficulty of factoring enormous numbers. For sufficiently large prime numbers, breaking RSA encryption entails factoring the modulus n into its prime factors, which becomes computationally impossible.

RSA has found pervasive use in various applications, such as secure communication, digital signatures, and protocols for key exchange. However, its security depends on adequately selecting the key size, safe key administration, and protection against implementation flaws and attacks.

#### **Advantages**

i. **Digital Signatures:** RSA facilitates the generation and validation of digital signatures, enabling secure message authentication and integrity verification. It allows the recipient to confirm the sender's identity and ensure the integrity of the message.

- ii. **Key Exchange:** RSA can be used for secure protocols like the Diffie-Hellman key exchange. It enables parties to establish a shared secret key over an insecure channel without surveillance or man-in-the-middle attacks being possible.
- iii. **Widely Supported:** RSA is extensively supported by cryptographic software and library collections. This algorithm has been implemented and evaluated in numerous programming languages and platforms.
- iv. **Efficiency for Short Messages:** Encryption and decryption operations performed with RSA are efficient for short communications. This makes RSA appropriate for protecting small quantities of data, such as symmetric encryption keys and digital signatures.
- v. **Compatibility:** RSA is compatible with various protocols and standards, allowing it to be interoperable across multiple systems and platforms. It is widely employed in secure email, SSL/TLS, VPNs, and other cryptographic programmes.

#### **Disadvantages**

- i. **Key duration:** The used key's duration affects RSA's security. As computing power grows, reduced key lengths become more vulnerable to attack. It is essential to employ sufficiently long keys to withstand future attacks. A key length of 2048 bits or greater is recommended for secure applications.
- ii. **Factoring Attacks:** The security of RSA is predicated on the difficulty of factoring large composite numbers. Nevertheless, developments in factoring algorithms, such as the General Number Field Sieve (GNFS), can threaten lower key sizes. Given sufficient computational resources, it is possible to factor RSA keys and decrypt encrypted data.
- iii. **Timing Attacks and Side Channels:** RSA implementations are susceptible to timing attacks and side channels. These attacks exploit information leaked during encryption or decryption, such as timing variations, battery consumption, and electromagnetic radiation. To mitigate these vulnerabilities, proper implementation and countermeasures are required.
- iv. **Padding Oracle Attacks:** RSA is susceptible to padding Oracle attacks when not correctly implemented. By repeatedly submitting altered ciphertexts and observing the server's response, attackers can exploit the padding oracle to recover the plaintext.

### **Real World Examples**

- i. Email Providers (Gmail, Outlook, proton mail, iCloud)
- ii. Web Browsers (Edge, Chrome, Opera, Brave)
- iii. VPNs (OpenVPN, SSTP protocol)
- iv. Chat Rooms and Messengers
- v. Peer-to-peer data transfer

# <u>Hash Algorithm – MD5(Message Digest 5)</u>

MD5 (Message Digest Algorithm 5) is a widely used cryptographic hash function that generates a 128-bit fixed-size hash value from an input message. Ronald Rivest created it in 1991 as an enhancement over previous MD hash functions.

The MD5 algorithm operates on data blocks and consists of multiple cycles of bit manipulation and modular arithmetic operations. The principal MD5 algorithm stages are as follows:

- i. The input message is padded to a multiple of 512 bits to ensure proper processing alignment.
- ii. The MD5 algorithm is supplied with a 128-bit initial value known as the initialization vector (IV).
- iii. Messages are divided into 512-bit blocks for processing, with each block undergoing a series of operations.
- iv. The MD5 algorithm applies bitwise logical functions, modular addition, and bit rotation in four iterations for each block. The block's bits are shuffled and altered to generate an intermediate hash value.
- v. After all blocks are processed, the intermediate hash values are combined to produce the final 128-bit value. A typical representation of the resulting hash is a 32-character hexadecimal string.

MD5's primary function is to generate a message's unique identifier or digital signature. The hash function is optimized for speed and efficiency, making it suitable for various applications, including data integrity checks, password storage (with added salt), and checksum verification.

MD5 is considered cryptographically flawed and insecure for specific applications. Different inputs can generate the same MD5 hash value, as discovered by researchers. Due to these flaws, MD5 is unsuitable for applications requiring robust collision resistance, such as digital signatures and certificate authorities.

# **Advantages**

- i. **Simplicity:** MD5 is a relatively straightforward algorithm, making it simple to comprehend and implement.
- ii. **Speed:** MD5 is computationally efficient and can generate hash values for large quantities of data rapidly.
- iii. **Compatibility:** MD5 hash values are widely supported and recognised, allowing for simple integration into various systems and protocols.

- iv. **Checksum Verification:** MD5 can be used to verify the integrity of files or data by comparing the hash values of the original data with those of the received data. The data has not been altered or corrupted if the hash values are identical.
- v. **Password Storage:** In the past, MD5 was frequently used to store passwords. It could convert passwords into hash values, enhancing security by not storing passwords in plaintext. Due to its vulnerabilities, safer password hashing algorithms such as bcrypt or Argon2 is now recommended.

# **Disadvantages**

- i. **Collision Attacks:** Different input messages can yield the same MD5 hash value, as demonstrated by practical collision attacks presented by researchers on MD5. This indicates that an attacker can construct two distinct inputs with the same MD5 hash, compromising the data's integrity.
- ii. **Preimage Attacks:** MD5 is susceptible to preimage attacks, in which an attacker can discover a message that generates a particular MD5 hash value. This indicates that an adversary can determine the original message based on its hash value, which violates the intended one-way property of a cryptographic hash function.
- iii. **Speed and Efficiency:** Although MD5's speed and efficiency were once viewed as advantages; they now contribute to its vulnerabilities. MD5 is more susceptible to bruteforce attacks, in which an attacker generates and tests many possible inputs to discover a collision or preimage due to its fast computation speed.
- iv. **Weak Security Strength:** By today's standards, MD5's 128-bit hash size is feeble, offering limited resistance to brute-force attacks. The advancements in computational power have drastically shortened the time required to locate collisions or preimages.
- v. **Cryptographic Breakdown:** The vulnerabilities in MD5 have resulted in its cryptographic collapse, rendering it unsuitable for security-critical applications. MD5's flaws have been exploited in real-world assaults, such as creating fraudulent digital certificates and malicious software with identical MD5 hashes to legitimate files.

### **Real World Examples**

- i. Verifying the integrity of files downloaded from the internet.
- ii. Storing passwords securely.
- iii. Digital signatures.
- iv. Verifying the authenticity of messages.

# <u>Implementation – AES</u>

#### **Problem Statement**

A file encryption and decryption program is developed using the AES-128 algorithm. The program provides a user-friendly interface to encrypt and decrypt files. The user can select the encryption or decryption mode, enter the file path, and provide a key for the encryption/decryption process. The program also validates the inputs and handle errors.

#### **Algorithm**

This code is written in python performs encryption and decryption of files which contain data in byte format using AES128.

The algorithm of the AES128 is:

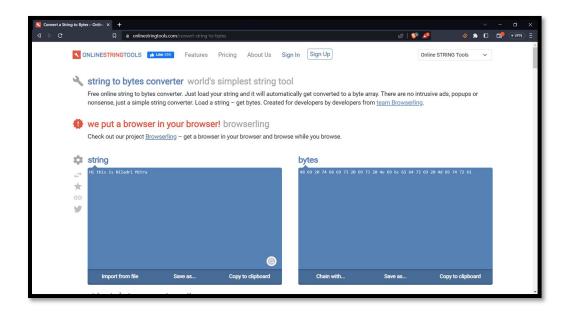
- i. Initialize variables:
  - a. Set nb as the number of columns in the state (4 for AES).
  - b. Set nr as the number of rounds in the cipher cycle (10 for AES with 128-bit key).
  - c. Set nk as the key length in 32-bit words (4 for AES with 128-bit key).
  - d. Define the hex\_symbols\_to\_int dictionary for converting hexadecimal symbols to integers.
  - e. Define the sbox list containing the AES S-box values.
  - f. Define the inv sbox list containing the inverse AES S-box values.
  - g. Define the roon list containing the round constant values for key expansion.
- ii. Define the encrypt function that takes input bytes and key as input.
  - a. Initialize the state array by dividing the input bytes into 4x4 matrix.
  - b. Generate the key schedule using the key expansion function.
  - c. Add the initial round key to the state using the add round key function.
- iii. Perform the main encryption rounds:
  - a. Iterate from round 1 to round nr-1.
  - b. Apply the sub bytes operation to the state.
  - c. Apply the shift rows operation to the state.
  - d. Apply the mix columns operation to the state.
  - e. Add the round key from the key schedule to the state using the add\_round\_key function.
- iv. Perform the final encryption round:
  - a. Apply the sub bytes operation to the state.
  - b. Apply the shift rows operation to the state.
  - c. Add the final round key to the state using the add round key function.
- v. Prepare the output:
  - a. Create an output list with 4 \* nb elements.
  - b. Flatten the state matrix and store the result in the output list.
- vi. Return the output list.

The algorithm of the main code which applies the above AES128 algorithm is:

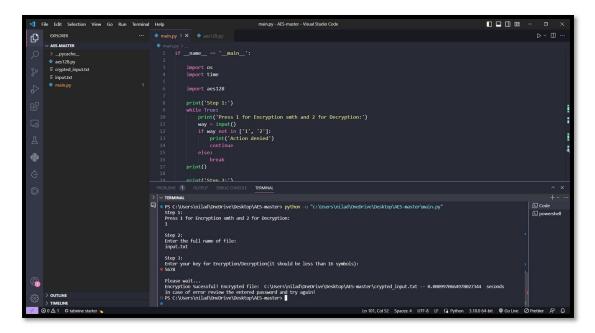
- i. Import the necessary modules: os, time, and aes128.
- ii. Enter a loop that continues until a valid input is provided.
  - a. Print "Press 1 for Encryption and 2 for Decryption:".
  - b. Read the input and assign it to the variable way.
  - c. If the input is not '1' or '2', print "Action denied" and continue to the next iteration.
  - d. Otherwise, break the loop.
- iii. Enter a loop that continues until a valid file path is entered.
  - a. Print "Enter the full name of the file:".
  - b. Read the input and assign it to the variable input path.
  - c. If the entered path is a file, break the loop.
  - d. Otherwise, print "This is not a file!" and continue to the next iteration.
- iv. Enter a loop that continues until a valid key is entered.
  - a. Print "Enter your key for Encryption/Decryption (it should be less than 16 symbols):".
  - b. Read the input and assign it to the variable key.
  - c. If the key length is greater than 16, print "Invalid Key! Use less than 16 symbols" and continue to the next iteration.
  - d. Iterate over each character in the key and check if it is a valid character (Latin alphabet or numbers).
    - If an invalid character is found, print "Invalid Key! Use only Latin alphabet and numbers" and continue to the next iteration.
  - e. Break the loop.
- v. Print a new line and "Please wait...".
- vi. Record the current time in the variable time before.
- vii. Read the input file data into the variable data using a binary read mode.
- viii. If way is equal to '1':
- ix. Initialize an empty list crypted data and an empty list temp.
- x. Iterate over each byte in data.
  - a. Append the byte to temp.
  - b. If the length of temp is equal to 16:
    - Encrypt temp using the aes128.encrypt function with the given key and assign the encrypted part to crypted part.
    - Extend crypted data with crypted part.
    - Clear temp.
- xi. If the length of temp is greater than 0 and less than 16:
  - a. Calculate the number of empty spaces needed to reach 16 by subtracting the length of temp from 16.
  - b. Iterate from 0 to empty spaces 1 and append 0 to temp.
  - c. Append 1 to temp.
  - d. Encrypt temp using the aes128.encrypt function with the given key and assign the encrypted part to crypted\_part.
  - e. Extend crypted data with crypted part.

- xii. Create the output path by joining the directory name of input\_path with a filename starting with "crypted\_" and the base name of input path.
- xiii. Open the output file in binary write mode and write the bytes of crypted\_data to the file.
- xiv. Otherwise, if way is equal to '2':
- xv. Initialize an empty list decrypted data and an empty list temp.
- xvi. Iterate over each byte in data.
  - a. Append the byte to temp.
  - b. If the length of temp is equal to 16:
    - Decrypt temp using the aes128.decrypt function with the given key and assign the decrypted part to decrypted part.
    - Extend decrypted\_data with decrypted\_part.
    - Clear temp.
- xvii. If the length of temp is greater than 0 and less than 16:
  - a. Calculate the number of empty spaces needed to reach 16 by subtracting the length of temp from 16.
  - b. Iterate from 0 to empty spaces 1 and append 0 to temp.
  - c. Append 1 to temp.
  - d. Encrypt temp using the aes128.encrypt function with the given key and assign the encrypted part to decrypted part.
  - e. Extend decrypted data with crypted part.
- xviii. Create the output path by joining the directory name of input\_path with a filename starting with "decrypted\_" and the base name of input\_path.
  - xix. Open the output file in binary write mode and write the bytes of decrypted\_data to the file.

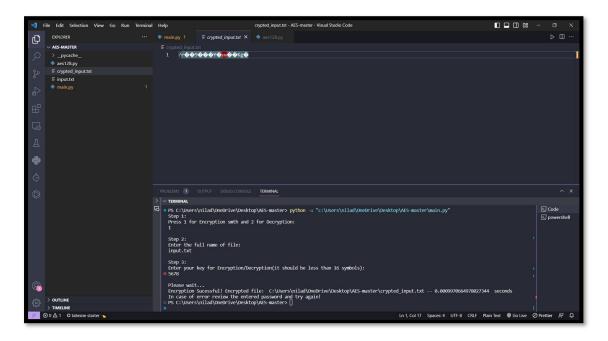
#### **Working Outputs**



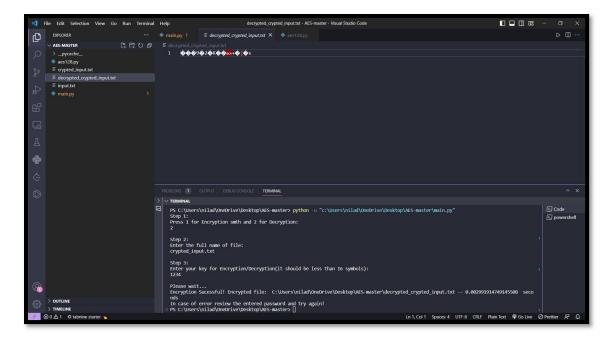
The byte form of a text is generated and stored in a file "input.txt".



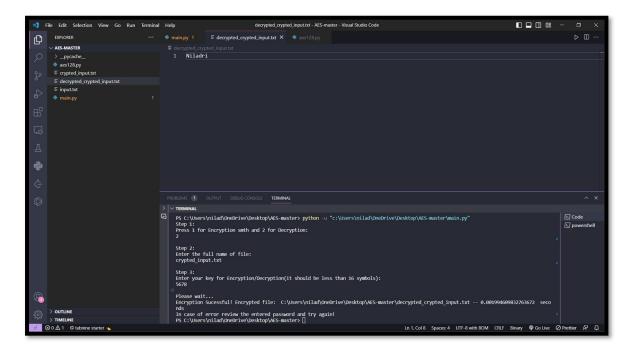
The main code is executed to encrypt the file.



This are the contents of the encrypted file 'crypted\_input.txt'.



On decrypting file with wrong key, the file shows corrupted contents.



On decrypting the file with correct key, the correct contents of the file are displayed.

# **Security Analysis**

The code has a few vulnerabilities:

- 1. **Input Validation:** The code lacks comprehensive input validation. While it checks some conditions, such as the user's input for encryption or decryption mode, the file existence check, and key length, it fails to validate other critical inputs thoroughly. For example, it does not validate the characters in the key or implement robust file path validation. This can lead to various security issues, including potential file system attacks or unauthorized access.
- 2. **Key Security:** The code does not implement strong key security measures. It allows keys longer than 16 symbols, but the AES-128 algorithm specifically requires a 16-byte (128-bit) key. Allowing longer keys may weaken the encryption strength. Additionally, the code does not enforce the use of strong, random keys or provide mechanisms for securely storing or transmitting keys. Strong key management is crucial for ensuring the confidentiality and integrity of encrypted data.
- 3. **Cryptographic Operations:** The code relies on the external module "aes128" for cryptographic operations. The security and integrity of this module are not verified within the code snippet provided. It is essential to use trusted and well-implemented cryptographic libraries to ensure the correctness and security of encryption and decryption processes.
- 4. **Padding:** The code implements padding for incomplete blocks of data using two different approaches. The first approach (padding v1) is commented out, and the second approach (padding v2) is used. However, the implementation of padding v2 is incorrect in the decryption section. It mistakenly uses the encryption function instead of the decryption function to process the padded block. This can result in incorrect decryption and potential data corruption or loss.
- 5. **Error Handling:** The code lacks robust error handling. While it displays some error messages for invalid inputs, it does not handle exceptions adequately or provide detailed error information. Proper error handling is crucial for preventing information leakage and maintaining the overall security and stability of the application.
- 6. **File Permissions:** The code does not consider or modify the file permissions of the output files it creates. This may result in insecure file permissions, allowing unauthorized access or modification of the encrypted or decrypted files.
- 7. **Side-Channel Attacks:** The code does not address potential side-channel attacks, such as timing attacks or power analysis. These attacks exploit information leaked during the execution of the cryptographic operations, potentially revealing sensitive information or weakening the security of the encryption algorithm.

To address these security concerns, the following measures can be considered:

- 1. Implement comprehensive input validation for all user-supplied inputs, including key validation and secure file path handling.
- 2. Enforce the use of strong, random keys of the correct length and implement secure key management practices.

- 3. Utilize trusted and well-implemented cryptographic libraries for encryption and decryption operations.
- 4. Ensure correct implementation of padding mechanisms for both encryption and decryption processes.
- 5. Implement robust error handling to handle exceptions gracefully and prevent information leakage.
- 6. Apply appropriate file permissions to the output files, limiting access to authorized users.
- 7. Consider and mitigate potential side-channel attacks through techniques such as constant-time implementations or countermeasures against timing and power analysis attacks.

# **Limitations and Challenges**

- 1. **Cryptographic Algorithm Selection**: One of the initial implementation challenges is selecting the appropriate cryptographic algorithm. In this case, AES-128 is chosen. However, there are various encryption algorithms available, each with its own strengths and weaknesses. Careful consideration needs to be given to factors such as encryption strength, performance, compatibility, and support when choosing a cryptographic algorithm.
- 2. **Integration of Cryptographic Libraries:** Integrating a cryptographic library, such as the external "aes128" module used in the code, can be complex. It requires understanding the library's APIs, documentation, and ensuring its compatibility with the programming language and platform being used. Proper integration involves correctly linking the library, managing dependencies, and handling any installation or configuration issues.
- 3. **Input Validation and Sanitization:** Implementing robust input validation and sanitization is crucial to prevent security vulnerabilities. It requires carefully checking user inputs, validating file paths, and performing proper sanitization to prevent common security risks such as path traversal attacks or injection attacks. Implementing a comprehensive and effective input validation mechanism can be challenging due to the various types of inputs and potential edge cases that need to be considered.
- 4. **Secure Key Management:** Implementing secure key management is a critical aspect of encryption. Generating strong, random keys, securely storing and transmitting them, and ensuring their confidentiality is challenging. Techniques such as key stretching, key rotation, and key storage in secure hardware modules may need to be considered depending on the specific requirements of the application.
- 5. Padding and Block Cipher Mode: Implementing padding for block cipher modes, such as the padding v2 mentioned in the code, can be complex. It requires understanding the specific requirements of the encryption algorithm and implementing the padding scheme correctly to ensure interoperability and security. Care must be taken to handle incomplete blocks, ensure proper padding validation during decryption, and prevent potential padding oracle attacks.

# **Conclusion**

In conclusion, cryptography plays a fundamental role in cybersecurity and ethical hacking by providing a strong foundation for protecting sensitive information and ensuring secure communication. It is an indispensable tool that enables confidentiality, integrity, authentication, and non-repudiation in various digital systems and applications.

The importance of cryptography in cybersecurity lies in its ability to safeguard data from unauthorized access, interception, and tampering. By employing encryption algorithms and techniques, cryptography transforms plain text into unintelligible ciphertext, making it extremely challenging for adversaries to decipher the information without the corresponding decryption key. This provides a layer of defense against data breaches, espionage, and unauthorized disclosure.

Ethical hackers, who are tasked with identifying vulnerabilities and weaknesses in systems, rely on cryptography to assess the security posture of organizations and uncover potential attack vectors. By attempting to decrypt encrypted data or exploit weaknesses in cryptographic implementations, ethical hackers can expose vulnerabilities and help organizations strengthen their defenses. This process highlights the critical role of cryptography in understanding the strengths and limitations of cryptographic systems and ensuring they are properly implemented and configured.