Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 12, No. 1, 2025

Discussion and Optimization of AES and SM4 Encryption Algorithms

Yibo Liu *

College of Computer Science and Technology, Inner Mongolia Normal University, Hohhot, Inner Mongolia Autonomous Region, 010010, China

*** Corresponding author Email:** lyb1968497372@163.com

Abstract: This paper creates a new direction for the operation optimization of Simplified Message Block Cipher for 128-bit Data encryption (SM4) national cipher algorithm, which combines with AES algorithm to provide a simpler encryption method for SM4 algorithm. The AES algorithm is for the whole file system-based f encryption scheme, and the implementation concept is to encrypt the file and maintain data security by encrypting files to encrypt and maintain data security. The advantages of this encryption method are that they have few key numbers and are easy to manage. The disadvantage is that when a user performs database operations, such as adding records, deleting records, changing records, and making selections, it is usually necessary to manipulate one or more records in a table. Currently, it's necessary for users to record the complete process of file encryption or decryption in a database, as the AES algorithm's sluggishness in encrypting the entire file encryption cipher block results in significant time consumption. SM4 algorithm, although the key generation is more cumbersome, but its encryption and decryption speed are relatively fast, so you can combine the two. This is a SM4 encryption scheme that maintains the security of the encryption scheme while having practical efficiency.

Keywords: AES Algorithm; SM4 Algorithm; Hybrid Encryption Algorithm; Group Encryption.

1. Introduction

With the development of science and technology, people's awareness of information security has become more and more important, and the way of cyber-attacks has evolved and upgraded. Cyber-attacks, occurring often, significantly affect societal stability, production, and people's lives. The rise in cyber threats to emerging technologies and scenarios is notable. The evolution of 5G technology has facilitated the shift from the Internet era to the Internet of Things period, necessitating increased data encryption speeds. Therefore, it is crucial to study how to realize the fast encryption of data while guaranteeing the security of data information in the Internet era.

The United States federal government has adopted the Advanced Encryption Standard (AES), a widely utilized symmetric encryption method, also referred to as Rijndael encryption in the field of cryptography, for block encryption. This block encryption method, embraced by the U.S. federal government as a substitute for DES, has undergone extensive analysis by various entities and enjoys global usage. Presently, conventional AES encryption methods restrict encryption speed and lead to substantial data wastage, attributed to the elevated complexity of encryption computations. The State Secret Algorithm (Simplified Message Block Cipher for 128-bit Data, SM4) is the first commercial cryptographic algorithm published by China's Cryptography Administration in 2006, which has played a great role in promoting the development of cryptography research in China [1].

Hybrid encryption algorithm is a novel encryption method that combines multiple types of encryption systems and algorithms, and its advantage is that it can flexibly choose different encryption mechanisms and algorithms for combination and construction according to the actual needs [2]. Scholars at home and abroad have carried out a lot of research on this method, Yang et al have studied the RSA encryption algorithm and DES encryption algorithm, which combines the characteristics of public key cryptography and symmetric cryptography to further improve the efficiency of encryption [3]. In terms of national cryptographic algorithms, Mo et al have also explored the high-performance hybrid encryption framework combining SM2, SM3, and SM4, enhancing the core hardware architecture [4]. This method addresses the drawbacks of individual cryptographic systems, including inadequate security, sluggish encryption rates, and complex key management. Abroad, Kuswaha et al. research on AES and RSA's combined security algorithm reveals that the suggested algorithm employs a dual-key approach, which enhances the resistance to linear attacks and enhances the security of encryption algorithm [5]. To sum up, hybrid encryption methods merge the benefits of various encryption techniques, enhancing not just the encryption algorithm's security but also considering the effectiveness of both encryption and decryption, unlike other single encryption methods that offer greater advantages. Drawing from the examination of existing hybrid encryption methods, this study merges the SM4 state secret algorithm with the AES algorithm, culminating in the creation of a hybrid encryption framework capable of executing rapid encryption and decryption on extensive information data sets while maintaining robust security. This aims to address the present demands for data encryption and decryption at both high speeds and security levels.

2. An Overview of Cryptography

In today's information society, information security has permeated all aspects of life, and recent cryptography research has its own characteristics. A cipher is simply a set of transformations E containing any parameter k. Assuming that the known message is m, the ciphertext c can be obtained by this transformation E, i.e. $c = E_k(m)$. This process is called encryption, and the arbitrary parameter k is the key.

After the encryption algorithm E is determined, because the key k is different, the ciphertext c is also different. However, not all transformations with arbitrary parameter k can be used as ciphers, which requires that the computation of $E_k(m)$ is not difficult, and that a third party, who does not have the key k, even if it captures the ciphertext c, cannot restore the message m, which is the plaintext from c. One of the two sides of the communication is called the sender and the other side becomes the receiver. The principle of traditional secure communication is shown in Figure 1.

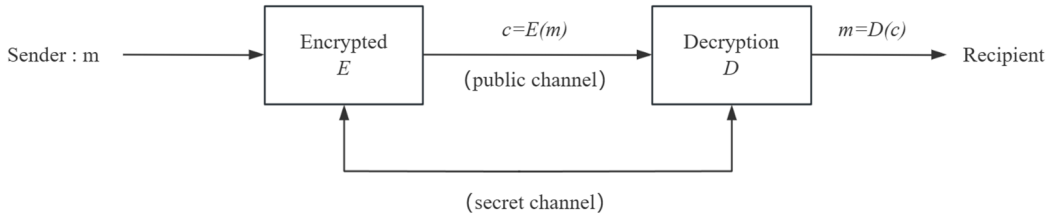

Figure 1. Schematic representation of conventional communication encryption

2.1. Keys

If cryptography is an important tool for securing information, the key is an indispensable core element in cryptography. In cryptography, a key is a string of data consisting of specific characters used to control the execution of encryption and decryption algorithms. The key can be regarded as a "password", and only the person who knows the key can correctly decode the encrypted information. The security of the key directly affects the security of the entire cryptographic system, therefore, a reasonable key management strategy is essential. The main role of the key is to protect the confidentiality, integrity and authentication of data in the encryption and decryption process. Keys can be categorized into symmetric keys, asymmetric keys and session keys. Methods involving symmetric key encryption employ an identical key for both encoding and decoding purposes. Common symmetric main algorithm include AES algorithm, DES algorithm, etc. The advantage of symmetric key is that encryption and decryption is fast, but the distribution and management of the key is its main challenge. Asymmetric key encryption algorithms use a pair of keys, a public key and a private key. The public key can be made public while the private key must be kept secret. Some of the common asymmetric key algorithms are RSA algorithm, ECC algorithm etc. The advantage of asymmetric key is that it simplifies the key distribution process, but its encryption and decryption speed is slower as compared to symmetric key. However, in some cases, in order to improve security, the system uses a temporarily generated key (session key) for one-time encryption and decryption. Commonly, this method combines public-key and secret-key encryption, starting with the encryption of the session key using the public key, which is then employed to encrypt the real data.

Managing keys is crucial for maintaining their security and efficiency. The primary components encompass generating keys, storing them, distributing them, updating, and revoking. Generating keys involves employing robust algorithms for generating random numbers to maintain the key's randomness and intricacy, thereby thwarting brute force decryption. Moreover, maintaining the keys in a safeguarded environment, such as a hardware security module (HSM) or a secure key management system, is crucial to prevent unauthorized access. Secure distribution of symmetric keys is a challenge and asymmetric key encryption is usually used to securely transmit symmetric keys. Regular key updates are an effective way to prevent security risks associated with long-term use. Once a key is compromised or no longer in use, it must be revoked promptly. Keys have a wide range of applications in a variety of information security areas, in protecting sensitive data in storage and transmission, and ensuring the confidentiality of information. In digital signatures the private key is used to sign the data and the public key can verify the authenticity of the signature, thus ensuring the integrity of the data and the authenticity of the source. In authentication process, key exchange protocols can confirm the identity of the parties to the communication, thus averting the intermediary attack. This method is extensively employed in online communication for encrypting key-based data transmission to guarantee protected connections.

2.2. Applications of Cryptography

Cryptography has a wide range of applications in modern society, covering a variety of aspects such as information protection, identity authentication, and data integrity. Cryptography plays a central role in network security. By encrypting communication, it ensures that information is not stolen or tampered with during transmission. For example, the HTTPS protocol is based on the HTTP protocol and utilizes SSL/TLS technology for secure encryption. In e-commerce, cryptography is used to protect users' personal information and payment information. The security and privacy of user transactions are ensured through secure payment protocols. Many businesses and organizations use cryptography to encrypt stored data to prevent data leakage and unauthorized access. The use of cryptography is especially important in industries such as healthcare and finance that require high data security. Blockchain technology uses cryptography to achieve decentralized storage and transparency of data, and transactions in the blockchain are verified through hash functions and digital signatures to ensure data immutability.

3. Common Encryption Methods

As the significance of information security escalates in the modern era, encryption emerges as a key tool for safeguarding data privacy and integrity. Encryption ensures the security of information during transmission and storage and prevents unauthorized access. Frequently used encryption techniques can be grouped into three categories: symmetric encryption, asymmetric encryption, and hash algorithms. The following will be introduced one by one.

3.1. Symmetric Encryption

Symmetric encryption employs a single key for both the encryption and decryption stages, making it a simple yet powerful technique for data protection. In this method, the sender encrypts the plaintext using a specific key, and the

receiver uses the same key to reverse the process and recover the original message. The primary benefit of symmetric encryption is its speed and efficiency, which makes it ideal for handling large amounts of data in a timely manner. This efficiency, combined with lower computational requirements, makes symmetric encryption highly suitable for applications that demand quick processing. Notable symmetric encryption algorithms, such as AES (Advanced Encryption Standard), are widely used due to their strong security and efficiency. AES stands out for its ability to provide robust protection through fixed-length keys (such as 128, 192, or 256 bits), while maintaining high performance across both software and hardware implementations. Consequently, AES has become a widely adopted encryption standard, safeguarding sensitive data in various industries and across numerous platforms. DES algorithm (Data Encryption Standard): DES is the early symmetric encryption standard, the key length of 56-bit, due to the short length of the key, easy to be violently decrypted, and has been gradually replaced by AES. 3DES algorithm (Data Encryption Standard): DES is the early symmetric encryption standard, key length of 56-bit. AES. 3DES algorithm (three-level data encryption standard): In order to solve the security problem of DES, 3DES encrypts data three times on the basis of DES, which improves security but its speed is relatively slow. The advantages of these symmetric encryption algorithms are fast encryption and decryption, simple algorithm implementation, and suitable for hardware acceleration. The disadvantages are the difficulty of key management, the security of the key directly affects the security of the whole system, once the key is leaked, all the data encrypted with the key will be at risk.

3.2. Asymmetric Encryption Algorithm

Known alternatively as public key encryption, asymmetric encryption employs a duo of keys: one public and one private. Encryption is performed using the public key, while the private key is employed for decryption. It's possible to disclose the public key, whereas the private key should remain confidential. The main advantage of asymmetric encryption is that it solves the key distribution problem. Common asymmetric encryption algorithms are: RSA algorithm: As the predominant public-key encryption method, RSA stands out for its high security, attributed to the intricate process of factoring in extensive numbers. Commonly used for safeguarding communication and electronic signatures, the ECC algorithm (Elliptic Curve Cryptography) utilizes elliptic curve mathematics to offer security comparable to RSA, albeit with reduced key lengths. Consequently, it finds extensive application in mobile devices and environments with limited resources. These asymmetric encryption methods offer benefits such as resolving key distribution issues, making the public key public without the necessity of transmitting the private key, and ensuring high security in scenarios involving large number decomposition and elliptic curve problems. The disadvantages are slow encryption and decryption speeds, complex algorithms and relatively difficult implementation.

3.3. Hash Functions

A hash function is a procedure that transforms input data of any size into a fixed-size output. It is widely used in areas such as data integrity checking and cryptographic storage. Hash function is characterized by irreversible, that is, from the output cannot be restored to the input. Common algorithms are: MD5 algorithm: although MD5 was once very popular, but due to its security issues (vulnerable to collision attacks), is now considered no longer secure. SHA algorithm (secure hash algorithm): The SHA-1 and SHA-2 series, encompassing SHA-256 and SHA-512, rank as some of the hash functions most prevalently utilized currently. SHA-2 is deemed to be more reliable compared to SHA-1. Bcrypt algorithm: a hash function designed for cryptographic storage, with adaptability, can increase the complexity of the hash calculation with the increase in computing power. Advantage is that the calculation speed is fast, suitable for data integrity verification. The principle of irreversibility guarantees the protection of data, preventing the restoration of initial data from the hash value. A drawback is the potential for security hazards due to misuse, such as employing MD5 for password storage. Collision attacks may result in different inputs generating the same hash value.

4. Principles of SM4 Algorithm

Functioning as a packet cipher, the SM4 algorithm features a 128-bit(16 bytes, 4 words) packet size and a corresponding 128-bit key length(16 bytes, 4 words). Its process of encryption and decryption employs a 32-round repetitive method, akin to that used in DES and AES, each round requires a round of keys (similar to DES, AES), the main operational structure to Z_2^{32} as the main unit symbol "$\oplus$ "represents a 32bit different-or operation, the symbol " <<<" denotes 32bit cyclic left shift i bit, "*MK*" denotes encryption key, " rk_i)" denotes round key, FK=(FK_0 ,FK_1 ,FK_2 ,FK_3) denotes system parameters, and (CK_0 ,CK_1 ,...,CK_{31}) denotes fixed parameters.

4.1. SM4 Encryption/Decryption Algorithm

Import: $(X_0, X_1, X_2, X_3) \in (Z_2^{32})^4; rk_0, rk_1, \ldots, rk_{31} \in Z_2^{32}$It's a wheel key.

Export: $(Y_0, Y_1, Y_2, Y_3) \in (Z_2^{32})^4$

1. $i = 0,1,2,\ldots,31$:

$$X_{i+4} = F(X_i, X_{i+1}, X_{i+2}, X_{i+3}, rk_i) = X_i \oplus T(X_i \oplus X_{i+1} \oplus X_{i+2} \oplus X_{i+3} \oplus rk_i)$$

2. $(Y_0, Y_1, Y_2, Y_3) = (X_{35}, X_{34}, X_{33}, X_{32})$

Where F is the wheel transformation function, T is the synthetic substitution, which contains the nonlinear transformation τ and linear transformation L , i.e., $T(.) = L(\tau(\cdot))$; τ has four parallel S-boxes inside, and L is the linear transformation. The decryption algorithm has the same operation flow as the encryption algorithm, except that the wheel key is used in the opposite order of encryption, and the synthetic substitution algorithm is shown in 4.2.

4.2. Synthetic Substitution T Algorithm

Import: $A = (a_0, a_1, a_2, a_3) \in (Z_2^8)^4$

Export: $C = (c_0, c_1, c_2, c_3) \in (Z_2^8)^4$

1. $B = (b_0, b_1, b_2, b_3) = \tau(A)$
$= (Sbox(a_0), Sbox(a_1), Sbox(a_2), Sbox(a_3))$
2. $C = L(B) = B \oplus (B <<< 2) \oplus (B <<< 10) \oplus (B <<< 18) \oplus (B <<< 24)$

4.3. Key Expansion Algorithm

Import: $MK = (MK_0, MK_1, MK_2, MK_3), MK_i \in Z_2^{32} (i = 0,1,2,3)$

Export: $rk_i \in Z_2^{32} (i = 0,1,2,\ldots,31)$

1. $(K_0, K_1, K_2, K_3) = (FK_0 \oplus MK_0, FK_1 \oplus MK_1, FK_2 \oplus MK_2, FK_3 \oplus MK_3)$

2. $i = 0,1,2,\ldots,31$:

$rk_i = K_{i+4} = K_i \oplus T'(K_{i+1} \oplus K_{i+2} \oplus K_{i+3} \oplus CK_i)$

T' The linear transformation in L' $:(B <<< 13) \oplus (B <<< 23)$, and the rest is the same as the T transformation in the encryption/decryption operation.

5. Principles of the AES Algorithm

AES (Advanced Encryption Standard) is a widely adopted symmetric encryption method used to protect data. It is based on the Rijndael algorithm and is known for its combination of strong security and high efficiency. AES supports three key lengths: 128-bit, 192-bit, and 256-bit. The longer the key length, the more secure it is, but the computational complexity increases. The longer the key length, the higher the security, but the computational complexity will also increase. The AES encryption algorithm primarily consists of four stages: byte substitution, row shifting, column mixing, and round key addition, while the AES decryption algorithm performs these steps in the reverse sequence.

5.1. AES Encryption/Decryption Algorithm

(1) Byte substitution: one-to-one substitution via AES-defined S-boxes.

(2) Row shift: when the key length is 128 bits, row 0 of the state matrix is shifted left by 0 bytes, row 1 is shifted left by 1 byte, row 2 is shifted left by 2 bytes, and row 3 is shifted left by 3 bytes.

(3) Column Mixing: the column mixing transformation is realized by multiplying the shifted state matrix with a fixed matrix to obtain the new state matrix after obfuscation, the specific process is shown in the equation in Fig:

$$\begin{bmatrix} s'_{0,0} & s'_{0,1} & s'_{0,2} & s'_{0,3} \\ s'_{1,0} & s'_{1,1} & s'_{1,2} & s'_{1,3} \\ s'_{2,0} & s'_{2,1} & s'_{2,2} & s'_{2,3} \\ s'_{3,0} & s'_{3,1} & s'_{3,2} & s'_{3,3} \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \begin{bmatrix} s_{0,0} & s_{0,1} & s_{0,2} & s_{0,3} \\ s_{1,0} & s_{1,1} & s_{1,2} & s_{1,3} \\ s_{2,0} & s_{2,1} & s_{2,2} & s_{2,3} \\ s_{3,0} & s_{3,1} & s_{3,2} & s_{3,3} \end{bmatrix}$$

Figure 2. Column Mixed Transform Equation Chart

(4) Wheel key addition involves a bitwise difference operation between the 128-bit wheel key Kj and the data in the state matrix. The method for generating the key Kj will be explained in the key expansion algorithm.

5.2. Key Expansion Algorithm

The primary key gets first placed in a 4x4 state matrix, with every column comprising 4 bytes, forming a word within each. These four columns contain words labeled Z[0], Z[1], Z[2], and Z [3], creating a word-wise array Z. Subsequently, this key array Z is extended by incorporating 40 new columns, culminating in a new key expansion array of 44 columns. The formation of new columns adheres to this recursive procedure:

(1) $Z[J] = Z[J-4] \oplus Z[J-1]$ (This formula applies when j is a multiple of four)

(2) $Z[J] = Z[J-4] \oplus T(Z[J-1])$ (This formula applies when j is not a multiple of four)

The function T, which is more complex, consists of the following three components:

a. Word Loop: Loop the four bytes in a word one byte to the left.

b. Byte substitution: The result of looping each byte in the word is substituted with an S-box lookup.

c. Cyclic constant dissimilarity: Dissimilarity operation between the result after byte substitution and the wheel constant Rcon[h] (where h denotes the current number of rounds).

6. Fast Dual-encryption Technique based on the AES and SM4 Algorithms

AES encryption algorithm, as a widely used symmetric encryption technology, has been generally recognized by the industry and is widely used for data encryption and protection. SM4 encryption algorithm is the national commercial cryptography standard in China, which adopts a 128-bit key and 128-bit data block, and is efficient and easy to implement, which is especially suitable for encryption needs in the Chinese market. Although AES is more efficient compared to some earlier algorithms (e.g., DES), the encryption and decryption process of AES may place a greater burden on computing resources and performance on certain low-performance devices, especially when dealing with large data volumes or high concurrency scenarios.SM4 is widely used in the government and financial sectors, although its relatively simple design makes it more susceptible to certain types of attacks compared to some complex algorithms (e.g., AES). types of attacks. Although the security of SM4 has not yet been breached, the simplicity of the design could mean potential weaknesses in the future.

The hybrid encryption algorithm combines the strengths of both the SM4 and AES encryption algorithms. In this approach, the SM4 algorithm is utilized to encrypt the actual data, while the AES algorithm is employed to securely encrypt the encryption key used by SM4. This dual-layer encryption mechanism not only enhances the overall security of the data but also simplifies key management by separating the encryption of the data and the key itself. This method ensures that the encryption key is securely protected, thus maintaining the confidentiality of both the data and the key. The detailed steps involved in the hybrid encryption process are as follows:

(1)The data requestor generates a pair of asymmetric public-private key pairs P and R and sends the public key P to the data owner.

(2)The data owner generates the key MK1, computes the hash value H1 of the plaintext M1, and subsequently uses the key MK1 to encrypt the data through the SM4 algorithm to generate the ciphertext Y. At the same time, the key MK1 is encrypted using the AES public key P to generate the key MK2.

(3)The data owner sends the ciphertext Y and the key MK2 to the data requestor.

(4)The data requestor decrypts the key MK2 by AES algorithm using the private key R and obtains the key

MK1.Then, it uses the key MK1 to decrypt the ciphertext Y by SM4 algorithm and recovers the plaintext M2.Next, it computes the hash value H2 of the plaintext M2 and compares it with H1. If H1 is the same as H2, the decrypted plaintext M1 is output; if it is different, the data is rejected.

7. Experimental Results and Analysis

It's vital to safeguard data encryption and decryption, as well as the processing speed, during the transmission of data to prevent unauthorized entry of confidential information. This study focuses on an in-depth assessment of the efficiency and safety of a novel hybrid encryption algorithm. To draw comparisons, we chose two renowned encryption methods—AES and SM4. In our research, we conduct an in-depth comparative examination of these algorithms, centering on their effectiveness regarding security and performance. Our focus is on analyzing the duration needed for encryption and decryption in the process of transmitting files, along with the level of security offered in both the encryption and decoding stages. The assessment offers crucial understanding of the hybrid algorithm's performance in practical scenarios, where strong security measures and rapid processing are crucial.

7.1. Encryption and decryption efficiency test

The AES, SM4, and hybrid encryption algorithms are employed to encrypt and decrypt a series of data volumes. For each algorithm, the time taken for both the encryption and decryption processes is recorded and analyzed. The resulting elapsed times for each algorithm are then summarized and compared to provide a clear understanding of their relative performance. A detailed overview of the results can be found in Table 1, which presents the time measurements for encryption and decryption across the different algorithms.

Table 1. Encryption and decryption average time statistics

data size	AES/s	SM4/s	Hybrid encryption algorithm/s
0.5G	5	2	3
1G	9	6	6
2G	22	10	13
5G	49	21	31
10G	78	39	51
20G	201	79	96
30G	278	125	155

7.2. Data Security Testing

Data security was evaluated using a brute force method. In the first step, the AES encryption algorithm, SM4 encryption algorithm, and the proposed hybrid encryption algorithm were employed to encrypt a 100KB data stream. This encrypted data was then transmitted over a public network using the FTP protocol. Next, the data in transit was intercepted using a Sniffer data interception tool, and the three different types of encrypted data were subjected to brute force decryption attempts. To assess both the efficiency and accuracy of the decryption process, the experiment was designed to perform six decoding repetitions within a 20-minute window. The results of these experiments are summarized in Table 2.

Table 2. Cracking statistics

data type	AES/session	SM4/session	Hybrid encryption algorithm/session
digital	5	4	2
text file	5	4	0
photograph	1	2	0

8. Conclusion

A fast hybrid encryption method based on AES and SM4 algorithms introduced and studied in this paper. By analyzing the advantages and disadvantages of both AES and SM4 encryption algorithms, a hybrid encryption scheme is designed to combine the advantages of AES encryption algorithm and SM4 encryption algorithm to improve data transmission security and processing efficiency. Experimental results show that the hybrid encryption algorithm can improve the encryption and decryption speeds greatly, and has better performance than the single algorithm. While this study demonstrates the potential of hybrid encryption algorithms in practical applications, there are still some improvements to be made. Computational complexity and key management in encryption process, for example, can be further optimized. In future, we will focus on how to improve the adaptability of the algorithm in different network environments and explore the combination with other encryption techniques to further improve the security and efficiency of the system.

References

[1] Lu S., Su B., Wang P., et al. (2016). Summary of the SM4 grouping cryptographic algorithm. Information Security Research, (11): 995-1007.

[2] Quan B. (2024). Application of hybrid encryption algorithm in computer network security. Electronic technology, (02): 184-185.

[3] Li W., Han, H. & Yang K. (2022). Combined encryption method based on AES encryption optimization algorithm. Information and Computer (theoretical edition), (15): 232-234.

[4] Li J., Mo Y., Su T., et al. (2022). Hardware design of high-speed hybrid encryption system based on the national secret algorithm SM2, SM3 and SM4. Research on Computer Application, (09): 2818-2825+2831.

[5] Kuswaha S., Waghmare S., Choudhary P. (2015) Data Transmission using AES-RSA Based Hybrid Security Algorithms. International Journal on Recent and Innovation Trends in Computing and Communication, 3(4): 1964-1969.