International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 20 (2019) pp. 3953-3958

Survey and Analysis for Achieving the Security of Data in Cloud

Manoj Tyagi
Research Scholar, Computer science, faculty of engineering
Mahatma Gandhi Chitrakoot Gramodaya Vishwavidyalaya, Chitrakoot, Madhya Pradesh, India.

Manish Manoria
Professor, Department of Computer Science and Engineering
Sagar Institute of Research and Technology, Bhopal, Madhya Pradesh, India.

Bharat Mishra
Associate Professor, Faculty of Science and Environment, Department of Physical Sciences
Mahatma Gandhi Chitrakoot Gramodaya Vishwavidyalaya, Chitrakoot, Madhya Pradesh, India.

Abstract

Cloud computing is on-demand access to computer system resources through the internet. Among various services provided by the cloud, most prominent is the cloud for storage. Because cloud services are present over the internet, which invites some trust issues among customers regarding integrity and confidentiality about the data, it is the challenge for the cloud to assure customers that their data is safe and can be retrieved whenever they want. This paper presents a survey of various schemes that applied to achieve confidentiality and integrity in the cloud system. Generally, encryption/decryption schemes and hashing techniques are used to obtain privacy and integrity, respectively. This paper also gives a comparative study of some encryption/decryption schemes as well as some hashing schemes.

Keywords: Cloud computing, encryption, decryption, confidentiality, integrity

INTRODUCTION

Cloud is a group of multiple servers, and cloud computing refers to allow to sharing of computer resources virtually. It also gives the facility of storing data at the cloud provided remote location via the internet. Cloud services model can be categorized into three models; IaaS, PaaS, and SaaS, where IaaS stands for Infrastructure as a service, PaaS stands for Platform as a service, and SaaS represents for software as a service. In IaaS, users can access all the computer resources like hardware, computer storage, and network virtualization from the cloud via the internet. It provides resources on sharing basis to the client without knowing any details about the location of the resources. The main advantage of IaaS is its location independence; anyone can access the computer resource irrespective of the location. Amazon web services (AWS) are famous IaaS providers [1]. PaaS provides tools that help clients for developing applications in the same environment. Lack of portability is the main issue in PaaS because it is difficult to transfer an developing tools from one platform to another. Google app engine is an accessible cloud that offers PaaS services. SaaS is delivering various user applications over the internet, where a developer provides an application on a server so it will be available for access to the customer. The application can only be accessed online because it is dependent on internet connectivity. Cloud computing has some models of cloud, such as the Public cloud model, Private cloud model, and Hybrid cloud model. The public cloud model can be accessible by all users. A private cloud model can only be available by a particular organization. The hybrid cloud model is the combination of the public cloud model and the private cloud model. Cloud offers the users to keep their data at the cloud and remove the storage overhead of users. Data service is available over the internet that creates a lot of security issues [2]. Data security is a significant issue that should be handled carefully; otherwise, it affects cloud popularity. Cloud computing faces many threats, then handling of these threats is required for survival of the cloud. Risks that can affect the cloud environment are an misuse of cloud services, insecure interface, privilege escalation, natural disasters, hardware failure, data breach or loss, malicious insiders, illegal access to the cloud system. If the cloud is secured from these threats, then cloud computing is an excellent option for all categories of users.

CLOUD SECURITY

Cloud computing is termed as the accessibility of computer assets via online. Its aim is to provide all computer system assets to their users as per their need over the internet. Cloud computing maintain quality service regarding availability, accessibility, flexibility, reliability to their clients. Third-party provider provides the service of cloud computing. This third-party provider owns the whole architecture of cloud computing, and cloud computing offers online services offered by IT technologies. Security is essential for making data confidential and safety. Security is critical in today's world of technologies. Cloud computing security is also termed as cloud security. It can be defined as a group of policies, applications, technologies, and controls that are used to ensure the safety of virtual IP, data, apps, services that are stored on the cloud [3]. In cloud security, cloud providers, as well as cloud consumers, both face security issues. The cloud computing companies or the providers should always take

care of the safety of the cloud infrastructure, data and applications of the client [4].

On the other hand, users should use strong passwords and authentication measures for the security of their data. There are some cloud computing controls to save the data from this type of incident. These are deterrent control, preventive control, detective control, and corrective control. Deterrent control is typically used to reduce the attacks that are being done in the cloud. A message sent to attackers by the system; if they proceed in the attack. The system gave a warning to the attacker. Preventive controls are used to increase the strength of the cloud. It is used to eliminate vulnerabilities. Detective controls are used to detect the attacks. If any incident occurs, the detective control detects it and informs the preventive or corrective control so that the issue can be addressed [5]. The curative control is used after the attack is done on the cloud or during the attack. It reduces the effects of the incident by limiting the damage. As it restore the system backups to rebuild the system. Now there are some security and privacy concerns of cloud computing. The first one is identity management. It means that the cloud providers should have their identity management methods to control unauthorized access to the resources and the information [6]. The second is physical security; the cloud service providers should take care of the hardware like the server, routers, and cables from unauthorized access. The third one is personal security. Now, after these, there are some uses of cloud computing. We usually do not notice it, but most of us use cloud computing services. We use online services for sending emails, watch movies, and edit documents. All these things are possible because of the cloud computing running behind it. Besides it, there are some more uses of it like create new apps and services, store, backup and recover data, host websites, and blogs, stream audio, and videos, etc.

Attacks on Cloud

Cloud computing is the availability of data storage and on-demand computer systems and resources. When it comes to storing data on the cloud, it involves sensitive information which should be kept safe from getting theft or being misused. Cloud is vulnerable to attacks, which is a significant threat to security. Some of the attacks are DOS, Side-channel, and authentication attacks.

DOS Attacks

It is an abbreviation used for denial of service attacks. In this attack, the hacker overloads a system by sending multiple requests at a time, which ultimately decreases the performance of the system and fails to respond to legitimate user's requests. These requests are sent intentionally to decline the efficiency of the server. The targeted system can be overloaded with junk files. These junk files occupy the network bandwidth. It makes the resources unavailable to the legal users waiting to use the system.

Channel Attacks

An electronic device is implemented in the vicinity of two systems, when they sharing data. After implementing the device, different behaviors of the two systems are recorded like the timings when the data is sent or received, or the radio frequencies emitted by the systems. These recorded values are studied, and the information is revealed. This study of these signals requires a high level of technical knowledge.

Authentication Attack

In cloud services, authentication is the point that is usually targeted by the attacker. Brute force is an authentication attack where all the possible combinations of the password are tried to break the security. Shoulder surfing where the activities of users are spied to get the password, phishing attacks where the attacker gets the password by redirecting the user to a fake website, key loggers are software programs that record each key pressed by the user.

Cryptography

The word cryptography means cipher writing. Cryptography is a way of protecting information at rest state and communication state through the uses of encryption, and only decryption key holder can access the information [7]. It has three process; key generation, encryption, and decryption process. Key generation process is responsible to create the keys. In private key cryptography, only one secret key is generated for both encryptions as well as decryption process. It is also called symmetric cryptography. Whereas in public key cryptography, two different keys are created; one is public key for encryption and other is secret key for decryption purpose. It is also called asymmetric cryptography. In cryptography, encryption is a procedure through which the plain text is changed into unintelligible form, known as ciphertext. A decryption is the reverse procedure of encryption which decodes the encrypted text into a readable form. Only authorized persons those having the decryption key can read and process the massage or information. After describing the secret and public keys now come towards on hash function which encodes the information. The hash function is used in authentication, key exchange and digital signature. The Hash function is also called one-way cryptography.

There are five main functions of cryptography nowadays. **Privacy or confidentiality**: In this, the sender conforms no one can read the information or message expects the receiver. Encryption/decryption is most popular for achieving it. DES, 3-DES, AES, RSA, ECIES is general techniques for maintain the confidentiality.

Authentication: A procedure of verifying users' identity. Generally username-password as a single factor is commonly used for verification purpose. But single factor verification is not fit for financial services and for sensitive data accessing. Then multifactor verification is required for secure access of such type of applications. Factors like OTP, biometric impression, smard-card are considered for next level verification. If considered two level then it is called two level authentication. If considered three factor then it is called three level authentication. OTP is globally used as second factor to verify the user, specially very useful in financial services and also in sensitive data accessing from cloud.

Integrity: A procedure which conforms that receiving message is original or not. Hashing techniques and massage authentication code approach are used for integrity purpose. Popular hashing schemes SHA-2 and SHA-3 are used in current time.

Non-repudiation: The method to provide evidence that the sender sends this message. Digital signature schemes are popular to achieve the Non-repudiation. RSA, DSA and ECDSA are generally utilized for creating digital signature. RSA takes more time to create a signature due to its key generation time. DSA is digital signature algorithm, also called Digital signature standard (DSS). ECDSA is based on elliptic curve for key generation then it is fast and efficient.

Key exchange: A method which securely shared the secret keys between the sender and receiver. Diffie-Hallman(DH) is popular scheme applied for key-exchange but it takes more time in key generation. Eliptic Curve Diffie-Hallman (ECDH) is taking less time compared to DH.

PRIVACY OR CONFIDENTIALITY

Generally, symmetric techniques and asymmetric techniques are applied to achieve confidentiality.

Symmetric Techniques

In symmetric techniques generally, DES, 3-DES, and AES are practically used in the past and also used in the current scenario.

Data Encryption algorithm (DES)

DES stands for the data encryption algorithm, which is block cipher of type of symmetric-key scheme published by the "National Institute of Standards and technology (NIST)" for cryptographic purpose. Its implementation is based on Feistel cipher and uses a 16 round Feistel structure. It uses 56 bits out of 64 since 8 of 64 bits were not used in encryption; they are used only as check bits. DES is the first algorithm of encryption and decryption, which was designed by a researcher at IBM in 1970 and is adopted by the government of the U.S. in 1977 for the encryption of commercial and defence information, and launch for public disclosure. Later it is used in day to day life, for example smart cart, sim-card and Set-top boxes [8].

DES Cipher Key weakness

As its key length is only 56 bit in the modern era of technology. From last two decades, machine was upgraded, and 56-bit key length can be easily broken by today's computer.

Brute Force Attack in DES

1992, 56-bit key length is too short as the turn to the new technology took place, and advance technology came system of 32 bit and more. Till 1990 it was a trusted algorithm for encryption. After that, some people raised the issue in this algorithm, but in 1998 Group of people (EFF) decrypted the encrypted code after the 56 hours; by latest speed machine and reduce the decryption time of 22 hours. As the DES uses a 56-bit key size to encrypted the message, it took a 2^56 attempt, which is equal to 72057594037927936 attempts to get the correct key.

3-DES (Three Data Encryption Standard)

3-DES is similar to DES encryption, but it repeats the same procedure of DES thrice, and the key length also changes. It uses 168 bit, which is thrice of the key length of DES. All those it is working well, but the time taken to encrypt the message and decrypted is much more than the DES encryption, which is about thrice. Because of time complexity; it cannot be used in regular use [9]. Then its successor came with name Advanced Encryption Standard, which is 128 bit and also efficient than the DES & 3-DES.

Advanced Encryption Standard (AES)

In 1997, NIST organized a competition for selecting the standard encryption techniques. Twenty-one algorithms participated in this competition. Based on their security and execution time efficiency, fifteen algorithms chosen in the first round, and five selected for the final round. These algorithms are Rijndael, MARS, RC6, Twofish, and Serpent. Finally, Rijndael was chosen as the AES, and it was selected due to Security, Cost, and implementation. AES is a symmetric-key block cipher, and hence, a single key is used for encryption and decryption both. This algorithm is to be used globally [10]. In AES, The block size of 128 bits and the key sizes are 128, 192, 256 bits. According to key size, the AES is of three versions – AES-128, AES-192, and AES-256. In AES-128, the key is of 128 bits and it haves 10 rounds. In AES-192, the key is of 192 bits and it haves 12 rounds. And in AES-256, there is the key of 256 bits, and it haves 14 number of rounds. In encryption/decryption, first, the Pre Round Transformation takes place, and then the 10, 12, or 14 rounds take place to depend on key size. Here every round except last contains four stages namely Sub-Bytes (substitution), Shift-Rows (permutations), Mix-Columns (mixing), Add-Round-Key (key adding)

Asymmetric Encryption

Whitfield Diffie and Martin Hellman, researchers at Stanford University, first publicly proposed asymmetric encryption in their 1977 in paper, "New Directions in Cryptography." The concept had been independently and covertly proposed by James Ellis several years earlier, while he was working for the Government Communications Headquarters (GCHQ), the British intelligence and security organization [11].

According to the researchers, asymmetric encryption has a pair of two keys. One is known as a public key, and other is known as a private key. The public key is used to encrypt the information, whereas the private key is used to decrypt the data. The receiver shares public key to the sender for encryption purpose, and the receiver can decrypt the data by a private key. The asymmetric encryption is also called public-key cryptography. Examples of asymmetric encryption techniques are ElGamal, RSA, DSA, ECC. Authentications are possible with asymmetric encryption. Asymmetric encryption is more secure than symmetric encryption due to secret key privacy. Asymmetric encryptions are very complex, and encryption/decryption is very time-consuming process. It is a deliberate process as compared to symmetric encryption; it is not appropriate for bulk message.

In 1978, Ron Rivest, Adi Shamir, and Leonard Adleman developed RSA cryptographic algorithm. This algorithm has two keys, one is the public key, and another is the private key. It is used in the encryption/decryption of data, authentication, and digital signature. It is based on the factorization of prime numbers, and its key generation and decryption process is slow and becomes slower with increasing key sizes. Then ECC is the alternative of RSA due to less time complexity of ECC as compared to RSA [12].

Elliptic Curve Integrated Encryption Schemes (ECIES)
ECIES is based on concept of elliptic curve. It is the complete security package which provides the confidentiality and integrity both. It uses the symmetric encryption so AES is used as symmetric function under ECIES in this paper. Message authentication code (MAC) is applied here to create a tag. It has three process; Key generation, encryption and decryption.
The elliptic curve for finite field is $z^2 = v^3 + av + b$
For binary field $z^2 + vz = v^3 + av^2 + b$

Key Generation process: It creates public key Q and private key d.

Process of Key Generation ECIES [12]
1. Choose $d \in_R [1, n-1]$
2. Compute $Q = dp$
3. Result (Q, d)

Encryption Process of ECIES
It converts plain text m into encoded form C with the help of AES scheme. First create keys k_1 and k_2 by using public key Q. Key k_1 is applied for encryption and k_2 is utilized for tag generation.

Process of Encryption ECIES [12]
1. Choose $k \in_R [1, n-1]$
2. Compute $R = kP$ and $Y = hkQ$
3. If $Y == \infty$ then
4. Go to Step 1.
5. End if
6. $(k_1, k_2) \leftarrow KDF(xY, R)$ Where xY , is the x-coordinate of Y.
7. Find $C = ENC_{k_1}(m)$ and $T = MAC_{k_2}(C)$
8. Output (R, C, T)

Decryption Process of ECIES

It changes encoded text C into plaintext m with the help of AES scheme.

Process of Decryption ECIES *[12]*
1. Execute confirmation for delimited Public key R
2. If the confirmation fails then refuse cipher text.
3. End if
4. Find $Y = hdL$
5. If $y == \infty$ then
6. Refuse cipher text.
7. End if
8. $(k_1, k_2) \leftarrow KDF(xY, R)$, x-coordinate of Y is xY.
9. Find $T' = MAC_{k_2}(C)$
10 if $(T' \neq T)$
11. Refuse cipher text.
12. End if
13.Find $m = DEC_{k_1}(C)$
14. Output (m)

First creates keys k_1 and k_2 by using secret key d. Key k_1 is utilized for decryption and k_2 is applied for tag generation. If tag is matched then decipher the encoded text otherwise refuse the ciphertext. In this way ECIES achieve the confidentiality, efficiency and integrity.

Integrity

Hashing is basic approach for achieving the integrity. A message digest is extensively used hash function, which provides the hash value. The message digest is a one-way function. Message digest takes variable size text and produces fixed-length message digest as output. This algorithm relies on the cryptographic hash function to initiate a distinctive value that is computed from data and a unique symmetric key [13]. These are also called encryption-only algorithm because the message-digest algorithm generates a value that is always used in an encrypted form only.

Message Digest Algorithm-2 (MD2)
In 1989, Ronald Rivest proposed this algorithm. The algorithm is streamlined for 8-bit computers. MD2 is designated in RFC 1319 [14]. Although MD2 is no longer appraised impregnable, even as of 2014, it remains in use in public key infrastructures as part of certificates generated with MD2 and RSA. The algorithm executes through a loop process where it permutes each byte. In 1997, Rogier and Chauvaud describe collisions of MD2's compression function [15]. In 2004, MD2 was manifested to be vulnerable to a preimage attack with time complexity equal to the 2^{114} application of the compression function. In 2009 the security updates were furnished debilitating MD2 in Open SSL and network security services.

Message Digest Algorithm-4 (MD4)
This algorithm was developed in 1990. The digest length is 128-bits. The algorithm has influenced to MD5, RIPEMD, and SHA-1 algorithm. The block size it uses is of 512 bits. It takes three rounds. It comprises of 48 operations, assembled in 3 rounds of 16 operations. The security of MD4 has been gravely compromised. A collision attack identified in 2007 can find a collision for full MD4 in less than two hash operations. MD4 is applied to find NTLM password-derived key digest on Microsoft windows NP, XP, Vista 7, 8, and 10.

Message Digest Algorithm-5 (MD5)
This algorithm is an extensively used hash function creating a 128-bit hash value. This hash function suffered from spacious vulnerabilities and achieve data integrity by identify international corruption. It was given in 1992 as RFC 1321. The flow of MD5 was exploited in the area, most famously by the FLAME MALWARE 2019. MD5 is more likely to length extension attacks. MD5 is still broadly utilized regardless of its well-documented feebleness and deprecation by security experts [16].

Secure Hash Algorithm (SHA)
A secure hash algorithm is a set of six hashing algorithms

introduced by NIST, which are applied for encoded any message. Hashing is the way of creating a fixed size code for input message. It has fixed length, which means that it is a combination of alphabets & numbers. A hash function, basically a mathematical function, is applied to the message to convert it into hash code.

The six different hashing algorithm's in the Secure Hash Algorithm (SHA) group are SHA-0, SHA-1, SHA-224, SHA-256, SHA-384 & SHA-512 [17,18]. All these six algorithms take variable-length messages in input and generate corresponding hash values of fixed-length in the form of output. The first four hash functions work on 512-bit message blocks split into 32-bit words, and the last two hash functions operate on 1024-bit message blocks split into 64-bit words. The working mechanism of all six algorithms is the same, as they all are an upgraded version of the previous one, overcoming the drawbacks of the earlier version.

Secure Hash Algorithm-1 (SHA1)

SHA-1 is the most popular hashing algorithm among all the 6 in the group. It takes an input message of maximum 264-bits and returns a 160-bits message digest as the output [17]. A message digest is the compressed form of the original message. This message digest is then given as an input to the Digital Signature Algorithm (DSA), the DSA then process the input message and generate a corresponding signature for that message in the output. In spite of being famous, SHA-1 is not very secure, as it is easy to understand, and it provides the same amount of security as provided by all it's competitor algorithms. It is not much complicated for the attacker to break this algorithm and conduct a brute force attack, preimage attack, or a dictionary attack.

Secure Hash Algorithm-2 (SHA2)

SHA-2 is the modified version of SHA-1, providing security against brute force attacks; these attacks are not much effective against SHA-2 as they are, against SHA-1. The SHA-2 also differs from SHA-1 in the message digest length, SHA-2 generates 224-bits or 256-bits digest [18].

Uses of SHA: They are generally used to encrypt the passwords that are going to be stored on the server. The servers only need to save the particular hash value of the users' actual password.

PERFORMANCE AND RESULT

Various hashing schemes and ECIES with AES are executed on 64-bit Intel-based Processor i5-8250 U with 8 GB. ECIES algorithm with AES-CBC mode is executed on five files from 1 to 5 MB, over 20 iterations. The results of various data are shown in figure 1. Execution time shows that encryption/ decryption time is approximately the same, and it is also faster than RSA. Various hashing schemes are also executed on 1 MB file over 20 iterations for determining tag generation time. A result shows that MD5 is more rapid than SHA schemes. The comparison is shown in figure 2.

CONCLUSION

Cloud computing attracts a lot of peoples and organizations through its revolutionary services. Mostly originations or peoples opts cloud storage services to keep their information to remote servers. The question is raised that how security can be ensured, where data keeps in outside remote servers because few security issues concerning the Authentication, Authorization, Availability, Integrity, Confidentiality, and Non-notoriety of information can possible.

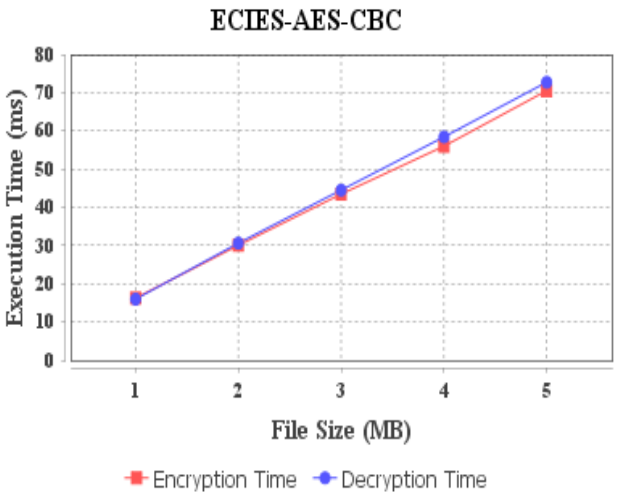

Figure 1: Execution time of ECIES-AES-CBC

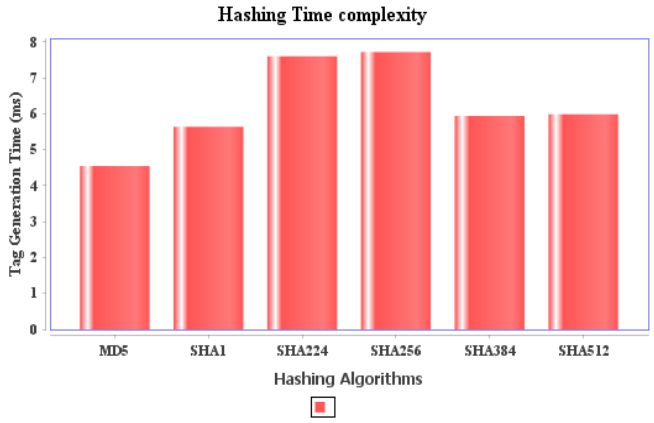

Figure 2: Tag Generation Time of Hashing schemes

This paper focused on confidentiality, integrity, and availability issues and their solution in a cloud environment. AES is the best-suited algorithm in the symmetric category because it proves its excellence over other algorithms while ECC is efficient than RSA regarding execution time. ECIES is a suitable option for achieving confidentiality and integrity. Moreover, it discussed hashing with implementation and finds that MD5 is faster than SHA algorithms, whereas the algorithm of the SHA family is more secure than MD5. ECDH is faster than DH under the key exchange protocol. ECDSA is more efficient than other digital signature schemes like RSA, DSA.

REFERENCES

[1] Buyyaa, R., Yeoa, C.S, Venugopala, S., Broberg, J., Brandic, I.: 2009 Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. J. Future Generation Computer Systems, Vol. 25, pp. 599-616, Elsevier

[2] Ashish Singh, Kakali Chatterjee "Cloud security issues and challenges: a survey", Journal of Network and Computer Applications, Elsevier , 2017, vol. 79, Pages 88-115

[3] Ali, M ,Khan, S and, Vasilakos, A (2015) "Security in cloud computing: Opportunities and challenges", Information Sciences, Vol.305, pp.357-383

[4] Ghorbel A, Ghorbel M, Jmaiel M,"Privacy in cloud computing environments: a survey and research challenges". J The Journal of Supercomputing 73(6): pp. 2763:2800, 2017

[5] EvanWheeler" Security Risk Reviews" Building an Information Security Risk Management Program from the Ground Up Security Risk Management 2011, Pages 239-257

[6] Rohitash Kumar Banyal, Pragya Jain , Vijendra Kumar Jain, "Multi-factor Authentication Framework for Cloud Computing" Fifth International Conference on Computational Intelligence, Modeling and Simulation , IEEE , 2013, Pages 105 – 110

[7] Tyagi M, Manoria M, Mishra B, (2017) "Effective Data Storage Security with Efficient Computing in Cloud" , In. conference CNC , Springer series Communications in

[8] Coppersmith D., "The Data Encryption Standard and its strength against attacks", J. IBM J. RES, DEVELOP. VOL. 38 NO. 3 MAY 1994

[9] Coppersmith D, Johnson D.B. , Matyas S.M., "A proposed mode for triple-DES encryption", J. IBM J. RES, DEVELOP. VOL. 40 NO, 2 MARCH 1996

[10] Nechvatal J, Barker E, Bassham L, Burr W, Dworkin M, Foti J, Roback E, " Report on the Development of the Advanced Encryption Standard (AES)", National Institute of Standards and Technology (2000)

[11] Diffie W, Hellman M E, "New Directions in cryptography" IEEE Transactions on Information Theory, Vol. IT 22, No. 6, pp 644-654, November 1976

[12] Martinez V G, Hernandez L, Dios A Q (2015), Security and Practical Considerations When Implementing the Elliptic Curve Integrated Encryption Scheme, Taylor and Francis, 39:244-269.

[13] Alkandari A A, Al-shaikhli I F, Alahmad M A (2013), "Cryptographic Hash Function: A High Level View" International Conference on Informatics and Creative Multimedia, IEEE computer and information science, Springer 2018, pp 153-164

[14] Kaliski B, The MD2 Message-Digest Algorithm RFC 1319 Network Working Group, April 1992.

[15] N. Rogier, Pascal Chauvaud, MD2 is not Secure without the Checksum Byte, Designs, Codes and Cryptography, 12(3), pp245–251, 1997.

[16] R. Rivest, "The MD5 Message Digest Algorithm", IETF RFC 1321, 1992.

[17] FIPS 180-1, Secure Hash Standard (SHS), National Institute of Standards and Technology, US Department of Commerce, WashingtonD. C.,1995.

[18] FIPS 180-2, Secure Hash Standard (SHS), National Institute of Standards and Technology, US Department of Commerce, WashingtonD. C.,2002.