RESEARCH ARTICLE OPEN ACCESS

Review On: Cryptographic Algorithms for Data Integrity Proofs in Cloud Storage

K.Devika[1], M.Jawahar[2]
Department of Computer Science and Engineering
K.S.R Institute for Engineering and Technology K.S.R Kalvi Nagar,
Tiruchengode Namakkal-637215,
Tamilnadu, India

ABSTRACT
In cloud computing , there are many cryptographic algorithm for storing and retrieving data , the cryptographic algorithm is more secure and highly well organized and successful in the field of authenticate information and keep the information private. This paper discusses cryptographic algorithms for data integrity proofs in cloud storage. Data integrity has been proved in cloud by enhancing the scheme known as proof of irretrievability (POR) using cryptographic algorithms. Here various cryptographic algorithms are compared. Those are different from one another based on its features. The important concerns that need to be addressed is to assure the customer of the integrity i.e. correctness of his data in the cloud. As the data is physically not accessible to the user the cloud should provide a way for the user to check if the integrity of his data is maintained or is compromised. In this paper a scheme provided which gives a proof of data integrity in the cloud which the customer can employ to check the correctness of his data in the cloud. This proof can be agreed upon by both the cloud and the customer and can be incorporated in the Service level agreement (SLA) and this scheme ensures that the storage at the client side is minimal which will be beneficial for thin clients.

Keywords:- Cloud Storage, Proof Of Retreivability (POR), Service Level Agreement(SLA), Data Integrity, Cryptographic Algorithm.

I. INTRODUCTION

Cloud storage moves the user's data to large data centres, which are remotely located, on which user does not have any control. However, this unique feature of the cloud posses many new security challenges which need to be clearly understood and resolved

The importance of cloud storage is reflected by the convenience of accessing the data from anywhere at any time through internet. For example, user store their data using gmail, ticket reservation etc. the data integrity to be proved by using the cryptographic algorithm. It use different cryptographic algorithm such as AES, DES, RSA,SHA,HASH .

II. RELATED WORK

Different studies are performing for proving data integrity, it based on file or data retrieval. It use Different cryptographic algorithm to retrieve the file or data.

A. AES Algorithm

The Advanced Encryption Standard (AES) is a symmetric-key algorithm approved by NIST which means same key is used for both encrypting and decrypting the data.

AES is based on the Rijndael cipher developed by two Belgian cryptographers, Joan Daemen and Vincent Rijmen Rijndael is a family of ciphers with different key and block sizes each with a

block size of 128 bits, but three different key lengths are used as 128, 192 and 256 bits. The input to the encryption and decryption algorithm is a single 128 bits block, this block is depicted as a 4x4 square matrix of bytes and copied into the State array, which is modified at each stage of encryption and decryption. After the final stage, State is copied to an output matrix. Similarly the key is depicted as a square matrix of bytes. This key is expanded into an array of key schedule words. In AES each word is four bytes, and the total key schedule is 44 words for the 128 bit key. The ordering of bytes within a matrix is by column. The cipher consists of number(N) of rounds depends on the key length as 10 rounds for a 16 bit key, 12 rounds for a 24 bit key, and 14 rounds for a 32 bit key. The first N-1 round consists of four transformation functions Known as SubBytes, ShiftRows, MixColumns, and AddRoundKey. The final round contains only three transformations, and there is a initial single transformation (AddRoundKey) before the first round. Only the AddRoundKey stage makes use of the key. Therefore the cipher begins and ends with an AddRoundKey. SubBytes is a non-linear substitution step where each byte is replaced with another according to a lookup table. ShiftRows is a transposition step where the last three rows of the state are shifted cyclically a certain number of steps. MixColumns is a mixing operation which operates on the columns of the state, combining the four bytes in each column.

AES algorithm is more reliable and secure and provides various key lengths to process the input files. By using this algorithm, the amount of control that an attacker can have on encrypted data depends on the encryption type and some specific details of some encryption modes can make the life a bit harder for the attacker if he wants to make surgical modifications.

B. DES Algorithm

Data Encryption Standard (DES) is a symmetric block cipher developed by IBM. This algorithm uses a 56-bit key to encipher/decipher a 64-bit block of data. The key is always presented as a 64-bit block, every 8th bit of which is ignored. However, it is usual to set each 8th bit so that each group of 8 bits has an odd number of bits set to 1.

The processing of plain text proceeds in three phases. First, the 64-bit plain text passes through an Initial Permutation (IP) that rearranges the bits to produce the permutated input. This is followed by the phase consisting of sixteen rounds of the same function, which involves both permutation and substitution functions. The output of the last (sixteenth) round consists of 64 bits that are a function of the input plain text and the key. The left and right halves of the output are swapped to produce the pre output. Finally the pre output is passed through a permutation $[IP^{-1}]$that is the inverse of the initial permutation function, to produce the 64 bit ciphertext. With the exception of the initial and final permutation, DES has the exact structure of a Fiestel cipher and it has key length as too short.

The substitution consists of a set of eight s boxes each of which accepts 6 bits as input and produce 4 bits as output. The key length argument goes like this. Assuming that the only feasible attack on DES is to try each key in turn until the

right one is found, then each capable of testing keys per second would find (on average) one key every 12 hours. Most reasonable people might find this rather comforting and a good measure of the strength of the algorithm.

C. RSA algorithm

RSA is a public key algorithm invented by Rivest, Shamir and Adleman. It uses different keys for encryption and decryption. Therefore it is called as asymmetric. RSA based on modular exponentiation. Numbers e, d and N are chosen with the property that if A is a number less than N, then (Ae mod N)d mod N = A. This means that we can encrypt A with e and decrypt using d. Conversely we can encrypt using d and decrypt using e.

• The pair of numbers (e,N) is known as the public key and can be published.
• The pair of numbers (d,N) is known as the private key and must be kept secret.

The number e is known as the public exponent, the number d is known as the private exponent, and N is known as the modulus. The concerns on the Key length, Anybody knowing the public key and they can use it to create encrypted messages, but only the owner of the secret key can decrypt them. Conversely the owner of the secret key can encrypt messages that can be decrypted by anybody with the public key. RSA also used for digital signature. Anybody successfully decrypting such messages can be sure that only the owner of the secret key could have encrypted them. This fact is the basis of the digital signature technique.

In digital signature technique, the RSA algorithm uses the public key to encrypt and the private key to decrypt in the encryption/decryption process. The private key is used to encrypt and the public key is used to decrypt in the digital signature. The sender using HASH algorithm to calculate the hash value of the file M, then generate the digital signature C from using the key to encrypt digital abstract and then M C together and sent to the receiver. The receiver receives the file M1 and digital signature C1, needs to verify that M and M1 are identical.

RSA is very slow with the problem of choosing the long keys when compared with a symmetric block cipher such as DES. The best solution is to use RSA for digital signatures and for protecting DES keys. When used in practice, RSA must be combined with some form of padding scheme, so that no values of M result in insecure cipher texts. RSA used without padding may have some problems. RSA also has a particular mathematical structure that can potentially be exploited without solving the RSA problem directly. To achieve the full strength of the RSA problem, an RSA-based cryptosystem must also use a padding scheme like Optimal Asymmetric Encryption Padding (OAEP), to protect against such structural problems in RSA.

D. SHA Algorithm

The Secure Hash Algorithm is the most widely used hash function developed by the National Institute of Standards and Technology (NIST) and published as a U.S Federal Information Processing Standard (FIPS). A prime motivation for the publication of the SHA was the Digital Signature Standard, in which it is incorporated. The four SHA algorithms are

structured differently and are named as SHA-0, SHA-1, SHA-2, andSHA-3.

SHA-0: This is the original version of Secure Hash Algorithm. SHA-0 produces the 160-bit hash function published in 1993 under the name "SHA". It was withdrawn shortly after publication due to an undisclosed "significant flaw" and replaced by the slightly revised version SHA-1.

SHA-1: A 160-bit hash function which resembles the earlier MD5 algorithm. This was designed by the National Security Agency (NSA) to be part of the Digital Signature Algorithm. In SHA-1, the cryptographic weaknesses were discovered and the standard was no longer approved for most cryptographic uses after 2010. SHA-1 is the most widely used of the existing SHA hash functions, and is employed in several widely used applications and protocols.

SHA-1 is very similar to SHA-0, but alters the original SHA hash specification to correct alleged weaknesses. It produces a 160-bit (20-byte) hash value that is typically rendered as a hexadecimal number, 40 digits long and differs from SHA-0 only by a single bitwise rotation in the message schedule of its compression function, this was done, according to the NSA, to correct a flaw in the original algorithm which reduced its cryptographic security. SHA-1 appears to provide greater resistance to attacks supporting the NSA's assertion that the change increased the security.

In Data integrity, Source control management systems such as Git and Mercurial use SHA-1 not for security but for ensuring that the data has not changed due to accidental corruption.

SHA-2: published in 2001, is significantly different from the SHA-1 hash function. A family of two similar hash functions, with different block sizes, known as SHA-256 and SHA-512. They differ in the word size. SHA-256 uses 32-bit words where SHA-512 uses 64-bit words. There are also truncated versions of each standard designed by the NSA known as SHA-224 and SHA-384. SHA-1 and SHA-2 are the secure hash algorithms required by law for use in certain U.S. Government applications, including use within other cryptographic algorithms and protocols, for the protection of sensitive unclassified information.

SHA-3: In 2012, NIST selected an additional algorithm, Keccak, for standardization under SHA-3. It supports the same hash lengths as SHA-2, and its internal structure differs significantly from the rest of the SHA family.

E. HASH Algorithm

A Hash algorithm is used for computing a condensed representation of a fixed length message/file. This is sometimes known as a message digest or a fingerprint. Some programs need a one way cryptographic hash algorithm, that is, a function that takes an arbitrary amount of data and generates a fixed-length number that hard for an attacker to invert, this stated as it is difficult for an attacker to create a different set of data to generate that same value. Hashing (or digest) algorithm should be computationally infeasible to construct a different message with the same digest.

The hash algorithm is a development of the MD4 algorithm invented by Ronald Rivest and

announced in 1990. Unfortunately, MD4 was flawed, so Rivest made some revisions, and the resulting algorithm was christened MD5. MD5 is a hashing algorithm that takes a message of up to 264 bits and reduces it to a digest of 128 bits (16 bytes). Instead of MD5 The new code to be written by using SHA-1. The SHA-256, SHA-384, or SHA-512 was used when more bits are needed in hash algorithm.

III. COMPARATIVE ANALYSIS

Algorithms used	Functions	Drawbacks
AES	i)SubBytes ii) ShiftRows iii)MixColumns iv)AddRoundKey	AES decryption cipher is not identical to the encryption cipher.
DES	i)Expansion D-box, ii)Whitener(XOR), iii)S-Boxes and iv)Straight D-Box.	i) Key length is too short. ii)Guarantees the integrity or authenticity (but not security of the message).
RSA	i)Encrypting messages, ii)Decrypting messages, iii)Padding schemes ivSigning messages.	i) Must use a padding scheme to protect against such structural problems ii) It has no random component. Therefore, an attacker can successfully launch a chosen plaintext attack against the cryptosystem.
SHA	Compression function	Functions are vulnerable to length-extension and partial-message collision attacks.
HASH	Message Digest	i)Worst behavior cause of excessive collisions ii)Results in poor performance

IV. CONCLUSION

This paper presents an extensive survey on data integrity proofs in cloud storage using cryptographic algorithm. Many new approaches are proposed in the field of data integrity proofs in cloud

storage. Many research issues have been highlighted and direction for future work has been suggested. Many open issues have been highlighted by the researchers such as dealing with data integrity proofs in cloud by different techniques future work will done.

REFERENCES

[1] Behrouz A. Forouzan, Debdeep Mukhopadhya, "Cryptography and Network Security" Second Edition,2011.

[2] William Stallings , "Cryptography and Network Security", Fifth Edition".

[3] Miles E. Smid and Dennis K. Branstad, "Data Encryption Standard past and furure" may-1998

[4] Dr Reinhard Wobst, "The Advanced Encryption Standard(AES): The Successor of DES" 2001.

[5] Douglas Stinson, Chapman & Hall/CRC, "Cryptography: Theory and Practice".

[6] Charles Kaufman et al, "Network Security: Private Communication in a public word".

[7] Dr. Prerna Mahajan and Abhishek Sachdeva, " A Study of encryption Algorithms AES,DES and RSA for security", 2013.

[8] Hamdan.O.Alan azi et al, " New comparative study between DES,3DES and AES within Nine Factors", 2010.

[9] Sunitra, "Comparative Analysis of AES and DES security Algorithms",2013.

[10] Krunal Suthar et al, "Analytical Comparison of Symmetric Encryption and Encoding Techniques for Cloud Environment", 2012.

[11] Cetin Kaya Koc , "Cryptographic Algorithms and Key Size Issues".