Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2021 doi:10.20944/preprints202111.0410.v1

Design and Implementation of Efficient Transmission of Cloud Data in Wireless Media

Virendra P. Nikam[1] and Dr. Sheetal S. Dhande[2]

***Abstract*— Nowadays, information security is a challenge especially when transmitted or shared in public clouds. Many of researchers have been proposed technique which fails to provide data integrity, security, authentication and another issue related to sensitivity data. The most common techniques were used to protect data during transmission on public cloud are cryptography, steganography, and compression. The proposed scheme suggests an entirely new approach for data security on public cloud. Authors have suggested an entirely new approach that completely makes secret data invisible behind carrier object and it is not been detected with the image performance parameters like PSNR, MSE, entropy and others. The details of results are explain in result section of paper. Proposed technique have better outcome than any other existing technique as a security mechanism on a public cloud. Primary focus of suggested approach is to minimize integrity loss of public storage data due to unrestricted access rights by uses. To improve reusability of carrier even after data concealed is really a challenging task and achieved through suggested approach.**

***Index Terms*— Data compression, data hiding, psnr, mse, virtual data, public cloud, quantization error**

I. Introduction

Public cloud storage is a place where multiple peoples (also called clients) store their data (ex. S3 - Simple Storage Service). Storage is called as public if it follow definitions mentioned below

1) Transparency
2) Availability
3) Security
4) Integrity
5) Authentication

Generally, public cloud is not considered as secure for storing sensitive data. Access protocols for public cloud data are very feasible which makes data mostly insecure. As it's name **public** implies, anyone can access data on the public cloud irrespective of it's identity. With access control protocol on public cloud restrict users for writing their sensitive data. A user of public cloud always threatens the misuse of data by unauthorized intruder on a cloud. The proposed concept is focused on how to maintain integrity, security, access control and reusability of data without 0% loss. Most of organizations prefer security of their data on a private cloud. A private cloud is owned by an organization itself and infrastructure is also maintained by an organization. It is too costly and not affordable in small scale industries. Public cloud is generally owned by third-party owners. Trusting on third party is sometimes highly risky. There are lots of rules available for accessing data on a cloud that restricts users from accessing data. The proposed concept will try to remove all those barriers that exist with available techniques so that a common person should able to use cloud storage data irrespective of his/her access control status.

A. *Public Cloud Properties*

1) ***Transparency****:* Client should not get exact location of cloud data from where it goes to/from cloud.

2) ***Availability****:* Client data should get Available $24 * 7$ Hours.

3) ***Security****:* Client data should get Stored encrypted or compressed or unreadable format on Cloud.

4) ***Authentication****:* Client data should get retrieved only to authenticate user.

5) ***Privacy****:* Client data should not exposed to unauthorized user and meaning should not open to others.

6) ***Integrity****:* Data should be automatically recovered after damage/corrupt.

Propose approaches is to combine data compression with virtual Data Hiding and stored result called as stego object on public cloud storage. The object on public cloud storage will be accessible to any personal with read, write and modify access. Below is the architecture of entire propose concept. Public cloud is not

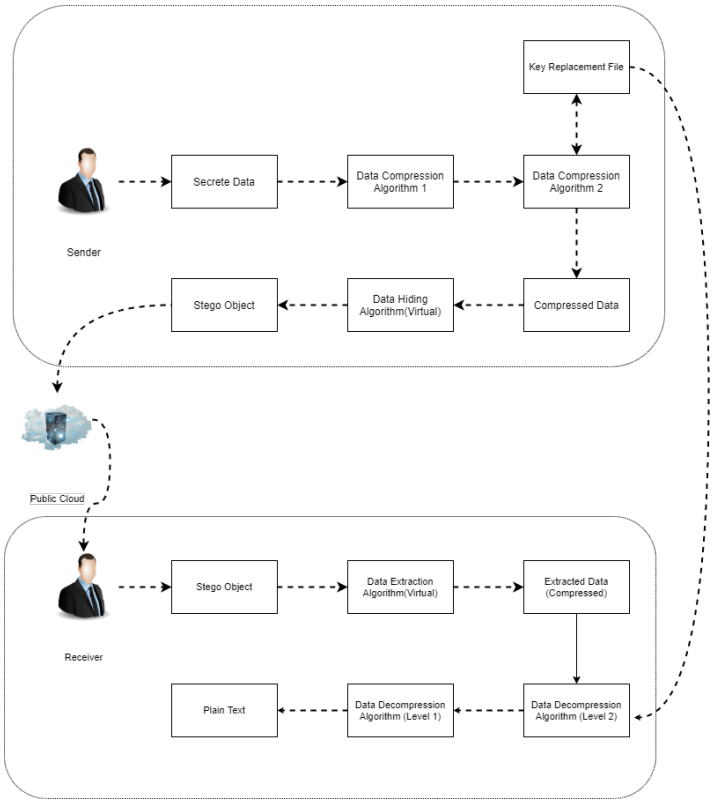


Fig. 1. Data Compression and Virtual data hiding

[1]Virendra P Nikam pursuing his PhD from sipna College of Engineering & Technology, Amravati. His main interest is in information and data security through cryptography, steganography and watermarking. He published more than 10 research paper in scopus indexed journals and conferences.

[2]Dr Sheetal S Dhande is a professor in department of Computer Science & Engineering at sipna College of Engineering and Technology, Amravati. Her major research work is on a database query Optimization and data security. She is having more than 20 years of teaching experience along with more than 5 years of research work experience at SGBAU Amravati. She guided more than 10 research scholar and also a member of board of study at SGBAU Amravati

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2021 doi:10.20944/preprints202111.0410.v1

a part of activity of propose concept. Sender initiate the data transmission process by selecting sensitive data which has to be transfer over insecure network. The data chosen by sender is first compress by converting it into consecutive bits of binary data. This binary data is fragmented into group of 8,7,6,....2 and its corresponding frequency of occurrence located. If fragment of data have the frequency of occurrence greater than or equal to 2, fragment string is replace with new character from key mapping table. Similarly when all the fragments with the length of 8 bits are done with replacement, replacement for next fragment of size 7 will start until it not reached to size 2. The process goes iterative for all the secret bits. The above compression is referred as the level 1 compression which is an import to level 2 and the process is repeated same as compression level 1. The output from compression level 2 consist of very less number of characters approximately equal to 2 to 3. It's a highest level compression. Any carrier object is enough to carry these two to three characters. These characters are hidden virtually behind the carrier object which maybe and image, audio, video or text. The virtual data hiding process generate stego object, it is stored on public storage(example s3 services). On public storage, many people's access the data with full rights. That's why there may have the possibility to corrupt or disturb the originality stego object on public storage. Receiver pickup stego object and extract compress data which is hidden virtually. The extracted data will be decompress with the help of key mapping file patterns which are shared either by sender or stored on public cloud. Entire process again require decompression of compress level2 data. The main intention of author is to protect sensitive data irrespective of access to stego object on public cloud. It may be happen that unknown person will destroy, update the object. But it will not affect hazards to sensitive data because it's not present over there. It's virtually hidden but extract a perfectly.

II. Literature Review

A. *Background History*

Data **compression** is one kind of cryptic technique where the size of data reduced. Data compression is effective when it is elastic in nature or sometimes called as lossless compression. Text compression mostly done with lossless technique whereas images or pictures compress with lossy compression technique. As text is a very sensitive format of data. Any error occurred due to compression or decompression will be easily identified by reader. The visual perception of text as a data is sensitive to user that makes compulsory to use lossless compression for text as a data. Whereas, images are less sensitive to its visual quality. Any small change in an image due to compression or decompression not easily identified by user which means lossy compression technique is possible on images. Joint photographic expert group (jpeg) is the best example of lossy compression or image. Proposed approach uses lossless compression which is elastic in nature with any size of compression data. Compression is the primary pillar of suggested approach that converts any length of data to minimum 2 to 3 characters compress data.

The process of hiding secret information behind carrier is called steganography. Steganography in real sense is an old concept which was first implemented 300 year before. In ancient time, peoples use this technique manually to transmit a message from one place to another with the help of fly. They remove hair on back neck of fly and then message is coded on their neck. They wait to grow hair and then transmit message to destination. Now a days, steganography comes in entirely different forms called as digital steganography where digital data get concealed behind digital carrier object. Carrier object maybe a picture, an audio or video. Different steganography techniques were introduced in recent 10 years that make data transmission more secure without inferring data by an Intruder. The first steganography techniques called as **Least Significant Bit substitution (LSB)** where a secret information bit get concealed at first (from right to left) position of carrier sample. To understand this, let's consider a carrier sample in binary format $(00010101)_2$ and secret bit is 0. After hiding secret bit 0, result carrier sample become $(00010100)_2$. This technique is simple to implement and preserves audiovisual perceptual quality of carrier. However due its simplicity, an intruder may easily locate secret bit position which create possibility of unauthorized data extraction. The drawbacks of LSB substitution technique were removed by hiding secret Bits in higher and higher LSB position. Moving from LSB to **Most Significant Bit (MSB)** increases quantization error Q_e. Increasing Q_e create a difference between result and original carrier object due to which it is very easy for an Intruder to locate an existence of secret information bit behind carrier. A Q_e varies from 0 to 255.

To minimize Q_e, approaches like spread spectrum analysis, wavelet analysis & sample selections approaches were suggested. All these approaches locate the best sample from carrier and then concealed secret information behind it. These methods reduce Q_e up to expected level. However, each technique has its own advantages as well as disadvantages. These approaches are time-consuming and complex to implement. None of the technique is ideal in terms of data hiding capacity, Q_e & time processing etc.

B. *Related Work*

M. Nosrati, R. Karimi, H. Nosrati, and A.Nosrati [1] in 2011 introduced a novel method that can find secret information behind 24 bits RGB color image. After hiding data, carrier samples are linked with each other so as to remember where secret information bit is kept. They are using a randomized sample selection approach for data hiding which may create complexity, difficulty in data extraction process. It uses what directional data link list to connect data hidden samples with each other. It also reserves some part of carrier sample to determine an existence of secret information bits.

Wen-Chung Kuo, Dong-Jin Jiang, Yu-Chih Huang [2] in 2008 proposed a block division-based data hiding and extraction method. This method divides carrier into multiple blocks and then histogram is generated for each. From histogram, minimum and maximum points are located that creates a space for secret data in each block that increase its data hiding capacity.

Soumyendu Das, Subhendu Das, Bijoy Bandyopadhyay,Sugata Sanyal [3] in 2008 introduced different data hiding techniques like image, audio, video & data steganography, etc. They focus on generating histogram of carrier to find out space for data hiding and cryptography to provide more security to sensitive data while transmitting it on insecure wireless network.

Ming Sun Fu and O.C. Au [4] in 2002 have proposed data hiding technique in halftone images. This technique has too high data hiding capacity when original multitone images are unavailable by force pair toggling. The visual perception of an quality of result image is so close with an original carrier that it can't be

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2021 doi:10.20944/preprints202111.0410.v1

possible to differentiate among them.

H. B. Kekre, Archana Athawale, Archana Athawale, Uttara Athawale [5] in 2010 proposed data hiding approach for hiding data in audios by generating stego audio carrier. They suggests not to hide data directly in LSB of an audio sample, but to calculate parity of sample first and then accordingly take decision to hide secret sample information bit. This approach creates an extremely difficult level for an intruder to guess where secret information bit is hidden in a carrier.

Xiaoyin Qi, Xiaoni Li, Mianshu Chen, Hexin Chen [6] in 2011 suggested data hiding approach in video as a carrier. This method uses add up approach to locate best sample from a carrier where secret information bit gets concealed. This technique functions by $4 * 4$ DCT block which is scanned in zigzag order to locate best matching sample.

T. Hong, W. Chen and H. Wu [7] in 2012 provided data hiding approach by splitting an image into multiple blocks. Data extraction is performed by finding smoothness of carrier sample. 4 borders of each block do not participate in data extraction process. The extraction process decreases when the size of a block is smaller. Extraction process on receiver side is performed by measuring noticeable smoothness to at-least one.

Mehdi Hussain, Ainuddin Wahid Abdul Wahab, Yamani Idna Bin Idris[15] in 2018 Did survey on various steganography methods in the spatial domain. the spatial domain is one where data bits are hidden directly in samples of carrier irrespective of modifying it by some operation. Many spatial techniques suggest hiding information either throughout an object or in a specific region of carrier. Both ways of sample selections i.e random and sequential are feasible in a spatial domain. Surrounding guard pixels provide additional security to sensitive secret bits while transmitting it over a wireless network.

Khan Muhammad, Jamil Ahmad, Seungmin Rho, Sung Wook Baik[16] in 2017 Suggest cryptographic approach for maintaining secret data security, integrity and authentication using steganography, Hue Saturation Intensity (HSI) model transformation, three-level Encryption Algorithm (TLEA) and Morton-scanning least significant bit substitution(LSB). it uses I- plane of an image to hide data using Morton scanning least significant bits and then encrypt this data bit using TLEA. In this approach, not all the samples are chosen for hiding data. The RGB model is converted into the HSI model and only one plane is used for Data Hiding. This means that only in 1/3 portion of an image, data get hidden. This approach selects samples sequentially from I- plane and once data bit hidden, selected samples kept in Dead Zone. Data Hiding capacity of this approach is limited due to the use of only one plane for data concealing.

G. Smitha, E. Baburaj [17] in 2016 Suggest data Hiding with Block based Edge Adaptive based on Least Significant Bit Matching Revisited (LSBMR). In this approach, boundary lines of different objects in a carrier are located and data is hidden base on the least significant bit substitution method. While hiding data, three directions from selected samples are considered that is top, bottom and right. Selecting sample 1 means by default selecting three samples for Data Hiding. This approach has much complexity and hence provide better security. It is an enhancement of an existing LSB Data Hiding technique. As data is hidden on boundary lines of different objects within an image, it's Data Hiding capacity is also very limited. Data Hiding capacity varies from image to image and from objects to objects within an image.

Jiantao Zhou, Weiwei Sun, Li Dong, Xianming Liu, Oscar C. Au Yuan Yan Tang [18] in 2015 proposed a Data Hiding approach in encrypted domain. A public key is used for image encryption and decryption hence no need to share a secret encryption key. This technique hides sensitive data in carrier and then encryption is performed with public-key cryptography. This technique avoids additional transmission cost request to transfer secret encryption scheme. On the receiver side, the image gets decrypted with different key and then hidden bits are extracted.

X. Zhang, Z. Qian, G. Feng, and Y. Ren [19] in 2014 proposes data Hiding in encrypted image base on lossless compression. The path of the encrypted image is compressed using the LDPC code and then embed additional secret data. The quality of encrypted data if considerable, on receiver side, hidden data get extracted successfully. Data hidden at 4th LSB of an image using LDPC code.

C. Qin, C. -C. Chang, Y.-H. Huang, and L.-T. Liao [20] in 2013 proposed a prediction-based reversible Data Hiding technique that select image carrier sample according to its distribution characteristics. An image in-painting technique is used to generate a predicted image which has similar structural content as that of original carrier image. Histogram of difference is shifted to embed secret information bit. On the receiver side, secret information bit can be extracted accurately without any change. Before to proceed for Data Hiding, the prediction process is conducted to estimate cover image pixels and Q_e.

C. *Gaps Reported*

Table I shows gap reported. Most of the authors or researchers have failed to identify threats, vulnerabilities and it's impact. On cloud storage access, users have to evaluate cloud service providers plan and need to analyze security status to protect their data from unauthorized access.

III. Problem Statement

A. *Unpredictable behavior*

User access data on public cloud, with almost all rights. If originality of file loss, it means an integrity of data losses. If user is authenticated, then no problem else wise it is highly unsecured. See table III for user Unpredictable behavior

B. *Firewall crash*

Generally, access to public cloud data is given either through user credentials or through a firewall. Depending totally on firewall for security is highly risky. If firewall crash, access to public cloud data is a easiest task. A firewall is only used to authenticate users. Once control passes form firewall, it don't have any control over user activity. A firewall is only used for purpose of authentication, not for entire security. Maintaining firewall is again costly and not affordable to small scale industries.

C. *Intruder attack*

An intruder is high skilled person or computer program that has a consistent watch on transmitting data over wireless media like air. It observes data on a network and tries to intercept it. Neither sender nor receiver has control on existence of intruder in a network. Once data get modified by an intruder, the receiver will get a wrong copy of data which may lead to wrong communication or failure of communication. Getting a wrong

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2021 doi:10.20944/preprints202111.0410.v1

TABLE I
GAPS REPORTED

Author Name	Functional Area	Gap Reported
Zhang et al. [9]	Cloud Environments, Security & Evaluation	Identify the threats and vulnerabilities and its impact of using the cloud.
Xie et al. [10]	Cloud Environments Security Evaluation	Users have to evaluate cloud service providers' plans, analyze the historical security status and, acquire the potential risk of the cloud service providers.
Tanimoto et al. [11]	cloud environment security analysis	Analyzed and extract risks of utilizing cloud computing by using the Risk Breakdown Structure (RBS) method. and using risk matrix method that classified risks into four kinds of risks.
Alhomidi and Reed [12]	cloud environment security analysis	Provided a mechanism for risk analysis using attack graph to present the relationships between vulnerabilities
Takabi et al. [13]	Based on Security Policies	Handle the security and trust issues of cloud computing environments by using various modules.
W. Zhao[14]	Based on Security Policies	Identifying the asset for the cloud deployment then evaluating risk for the asset.

TABLE II
CLOUD ACCESS MAPPING

Deployment Model	Managed By	Infrastructure Owned By	Infrastructure Located At	Accessible and Consumed By
Public	Third party provider	Third party provider	Off-premise	Untrusted
Private	Organization	Organization	On-premise	Trusted
			Off-premise	
	Third party provider	Third party provider	On-premise	
			Off-premise	
Managed	Third party provider	Third party provider	On-premise	Trusted or Untrusted
Hybrid	Both organization	Both organization	Both on-premise and off-premise	Trusted or Untrusted
	and third-party provider	and third-party provider		

TABLE III
BEHAVIOR OF CLOUD USERS

Operation	Loss	Change
Update	Recovery Not Possible, Authenticate, Non- Authenticate	Change Object
Delete	Recovery Not Possible, Authenticate, Non- Authenticate	Change Object
Read	Authenticate, Non- Authenticate	No Change
Write	Authenticate, Non- Authenticate	Change Object

copy of data on a network means that the network is highly unsecured for sensitive or secret data. Intruder may also misuse data on a network.

D. *Third party dependency*

Third-party is responsible for all activities of public cloud data management. Third-party service is highly costly and generally not affordable by small scale industries. Depending on only third party is highly risky because, collapse of third party may lead to collapse of entire public cloud storage management.

E. *Multiple access*

When non-owner accesses files on public cloud, the privacy of file gets violated. Restriction of file access only to owner, minimizes reusability of data on cloud. Accessing file by nonowner means an losing an integrity which means destroying hidden information in file.

F. *Backup crashed*

If database and backup server of a public cloud storage crash, then it leads to collapse of entire public cloud data management. With collapse of backup server and database, it's not possible to revert to original state.

The aim of author is to overcome all these problem statements reported in public cloud domain. Virtual data hiding is the best solutions to overcome these problems which is not possible without enhance data compression technique.

IV. PROPOSE METHODOLOGY

Figure 1 represents to entire flow of propose approach. This section represents detail view of each block flow.

A. *Architecture view - Sender*

The architecture view of sender side is shown in figure 2. The sender selects a secret information file which is given as an input to compression algorithm and generates compressed key file with compressed data. A compressed key file is a set of compressed patterns for each compressed word. A compressed data is given as an input to data hiding algorithm which generates stego carrier. Data hiding mechanism can be implemented with any approach like MSB, LSB, spread spectrum, and others. The carrier object may be an image, audio, video, or text. A suggested approach is flexible with any kind of steganography technique. Once the stego carrier object is generated, it is transmitted through transmission media to receiver. Algorithm I explained the processing steps on sender side. Figure 3 indicate the transmission of compressed key pattern file to authenticated user. However, there is an alternative exist i.e. Store compress key pattern file on public cloud. Generated stego file transmitted over wireless media like air and stored in public cloud storage. Public cloud storage is

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2021 doi:10.20944/preprints202111.0410.v1

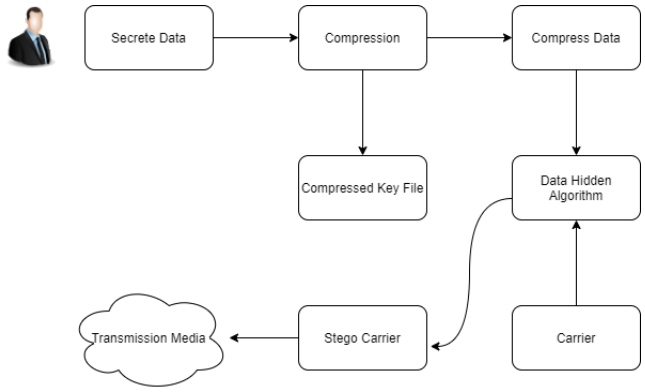


Fig. 2. Architecture view - Sender

Algorithm 1 Processing Steps - Sender

Require: Secrete file, Carrier

1) Start
2) Input secret data to compressed pattern.
3) Generate compressed data and compressed key file.
4) Input compressed text to data hiding algorithm which concealed into carrier object.
5) Stego carrier and compressed key file are transmitted over wireless media or stored on cloud
6) Stop

not a part of activity of either sender or receiver. Once's file is stored on cloud, it is available for open access. Figure 4 shows the entire picture after file available on cloud storage. N number of users will access file without any permission by author

*1) **Data Compression-Training**:* Data compression is extremely vital steps in proposed algorithm. If output of compress algorithm have result character approximately closer to 1, it means algorithm is good else algorithm needs focus more. Algorithm 2 represents data compression based on key patterns for each and every word. It's a self adaptive learning approach where if pattern is not present then it adds a new pattern in compress key file. Every time, compress key pattern file increases in size for new secret data.

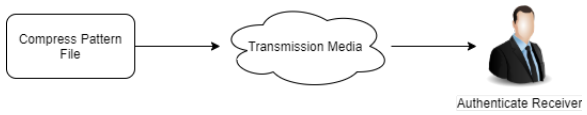


Fig. 3. Transmission of Compress Pattern file to authenticate user)

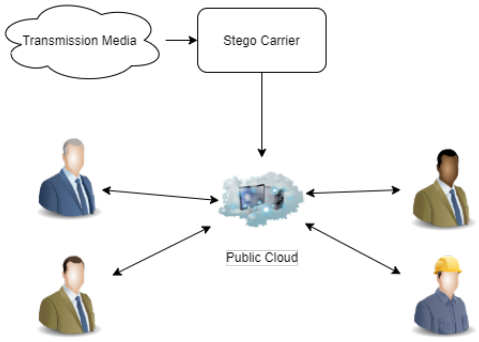


Fig. 4. Transmission of Stego file to authenticate user

Algorithm 2 Data Compression-Training

Require: Secrete Data Bits, Compress Key Patterns

1) Start
2) Load Compress Text (CT) – Level 1
3) Enter Threshold (T_h)
4) Initialize Char []RC= new char[] 'a','b','c','d','e','f','g','h','i','j','k','l','m','n','o', 'p','q','r','s','t','u','v', 'w','x','y','z','A','B','C','D','E','F','G','H','I','J','K', 'L','M','N','O','P','Q','R', 'S','T','U','V','W','X','Y','Z'; Count=0;
5) Initialize Net (Network File – Training File);
6) For I=0: length (CT)
   Sub-Pattern = Extract SP (CT, I, T_h)
   If (FoC(SP) >=2)
   RP = "-" + Length (SP) + RC[count] + Position_of_SP_in_ Net+ "- ";
   CT = Replace (SP, RP);
   Add RP to Net (Network File – Training File);
   Count ++;
   End
   End
7) Save CT
8) Stop

*2) **Data Compression-Testing**:* The only difference between data compression training and testing is the presence of compress key file pattern. In training, algorithm adds new pattern to compress key file whereas in testing, an existing patterns are utilized for data compression and no new pattern added to compressed key file. From complexity point of view, time required for completion in training phase is more than compression in testing phase. Algorithm 3 explain compression testing

Algorithm 3 Compression - Testing

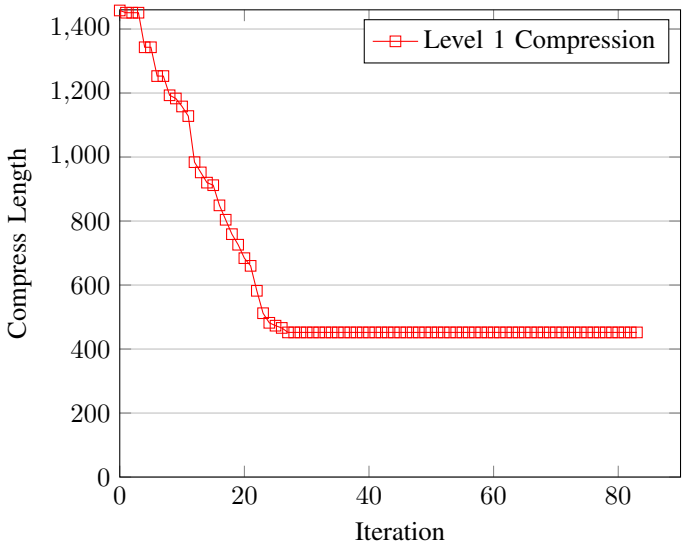
Require: Secrete Data, Compress Key Patterns

1) Start
2) Load Compress Text (CT) – Level 1
3) Enter Threshold (Th)
4) Load Net (Network File – Training File);
5) For I=0: length (CT)
   Sub-Pattern = Extract SP (CT, I, Th)
   If (SP_Found_in_Net == True)
   CT = Replace (SP, Pattern_in_Net);
   End
   End
6) Save CT
7) Stop

*3) **Data Compression - Level 1**:* Level 1 data compression is the first step of compression where frequency of occurrence of pattern is located in secret information bits. As per equation 1 to 11, if frequency of occurrence of pattern greater than or equal to 2, then pattern will be replaced with single character. If replacing character is present in compressed key file pattern, then it replace with existing pattern else next character from replacing pattern selected for replacement. The process continues until all patterns of secret information's bits not processed. The

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2021 doi:10.20944/preprints202111.0410.v1

amount of time required for processing for compression level 1 is directly proportional to amount of secret information bits and patterns in compressed key file. Graph represents output of compression level 1 which indicates that compression goes stable after specific iteration. Compression becomes stable because it reach to it's minimum sub pattern length.



$$f(x) = f(x) - (c * f(p) - c) \tag{1}$$

Where
f(x) = Compress Text_Length, c = Occurrence Frequency, f(p) = Sub_Pattern_Length

$$\int_a^c f(x) = \int_a^b f(x) + \int_b^c f(x) \tag{2}$$

$$= \int_a^b (f(x) - (c * f(p) - c) + \int_b^c (f(x) - (f(p) - 1))) \tag{3}$$

$$= \int_a^b f(x) - \int_a^b c * (f(p) - 1) + \int_b^c f(x) - \int_b^c f(p) + 0 \tag{4}$$

$$= \int_a^b f(x) - (\int_a^b c * f(p) - \int_a^b c) + \int_b^c f(x) - \int_b^c f(p) \tag{5}$$

$$= \int_a^b f(x) - (\int_a^b c * \int_a^b f(p)) + \int_a^b c + \int_b^c f(x) - \int_b^c f(p) \tag{6}$$

$$f(x) = \{n \approx 1 \rightarrow Good \quad \} \tag{7}$$

$$f(x) = \{n \neq 1 \rightarrow bad \quad \} \tag{8}$$

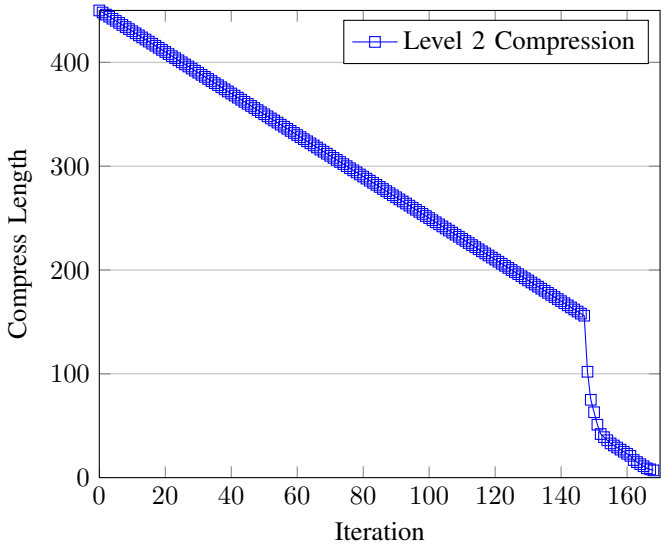
Smallest change in Secrete Binary with respect to x

$$\Delta x = \frac{\delta f(x)}{\delta x} = \frac{\delta (f(x) - (c * f(p) - c))}{\delta x} \tag{9}$$

$$= 1 - \frac{c(f(p) - 1)}{\delta x} \tag{10}$$

$$= 1 - 0 = 1 \tag{11}$$

*4) **Data Compression - Level 2**:* Level 2 compression is performed with compressed key file pattern. After level 1 compression , compress patterns are further fragmented into different sub patterns and then it's corresponding replacement character is search in compressed key file. If pattern present in compress key file, then it is replace with replacement character. The output of compression level 2 is almost closer to character 1. Below represent to compression level 2 output. If an output of compression level 2 decreases from n^{th} characters to 1 character, then it is considered as good compression else, compression algorithm needs modification. An equation from 12 to 18 represents compression level 2 processing steps. The minimum change in secret binary data is approximately equal to one character (as per equation 18).
Level 2 compression is performed on the output of Level 1 compression text. The repetitive string pattern in compression Level 1 output stream is located and replace with new characters from key replacement file. From graph, it is clear that the compression at level 2 inversely decreases with number of iterations and then exponentially reaches to its minimum value. The effectiveness of output of compression level 2 is totally depend on patterns available in key replacement file. Total numbers of key replacement files are 7.



Level 2 Text Compress with

$$f(y) = \int_a^b (f(-y) + m) + \int_b^c e^{-y+m} \tag{12}$$

$$= -\int_a^b f(y) + my + \int_b^c e^{(m-y)} \tag{13}$$

$$= -\int_a^b f(y) + my - (e^{(m-y)}) + m \tag{14}$$

Smallest change in Secrete Binary with respect to y

$$\Delta x = \frac{\delta f(y)}{\delta y} = \frac{\delta (f(-y) + m)}{\delta y} + \frac{\delta (e^{(-y+m)})}{\delta y} \tag{15}$$

$$\Delta x = \frac{\delta f(y)}{\delta y} = -\frac{\delta y}{\delta y} + \frac{\delta m}{\delta y} + e^{(-y+m)*1} \tag{16}$$

$$\Delta x = \frac{\delta f(y)}{\delta y} = -1 + 0 + e^{(-y+m)} \tag{17}$$

$$\Delta x = \frac{\delta f(y)}{\delta y} = -1 + e^{(-y+m)} \tag{18}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2021 doi:10.20944/preprints202111.0410.v1

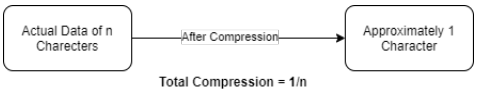


Fig. 5. Character Compression

*5) **Character Level Compression**:* The proposed concept will compress data up to $\frac{1}{n}$ character. Where n = total input characters. However, authors are expecting this output in maximum cases. Competitions will entirely depend on how compressed key patterns get trained and the length of replacing patterns. Hence

Key Replacement File	Pattern Size
8	8
7	7
6	6
5	5
4	4
3	3
2	2

TABLE IV
KEY REPLACEMENT FILE

the total patterns in key replacement files are $8! + 7! + 6! + 5! + 4! + 3! + 2! = 46,232$

*6) **Data hiding**:* Propose concept is flexible with any existing data hiding algorithm. It's performance is completely depend on how the data compression algorithm performs?. Algorithm 4 represents data hiding process where RGB components from carrier object (image) are extracted. Each extracted RGB components are logically AND with secret character in binary format. The logical XOR is chosen because of it's reversibility during data extraction.

Algorithm 4 Data Hiding

Require: Carrier Object, ROI, Secrete Data Bits

1) Start
2) Input Carrier (Cr).
3) Extract Features (In Case of Image as a Carrier A, R, G, B).
4) Load CT
5) Initialize Count=0;
6) For I=0: Width (Cr)
   For J=0: Height (Cr)
   Replace A (I, J) = CT[Count] XOR 255;
   Pixels (I, J) = Color(A, R, G,B);
   End
   End
7) Stop

B. Architecture view - Receiver

Figure 6 represents an activities performed by receiver. If receiver is authenticated, then he/she will selects stego carrier object from public cloud as well as a compressed key file. Selected carrier file is given as an input to data extraction algorithm which extracts data in compressed format. This compressed data is decompressed with help of a compressed key pattern file. Decompressed data is in plain text readable format which is exactly similar to data hidden by sender during process of data hiding. Algorithm 5 explained processing steps by receiver.

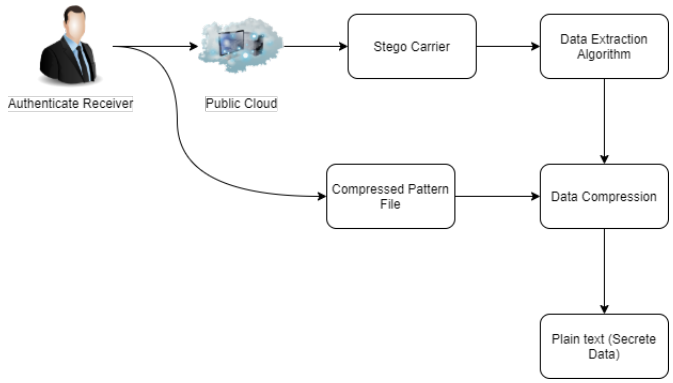


Fig. 6. Architecture view - Receiver

Algorithm 5 Processing Steps - Receiver Side

Require: Stego Carrier, Compressed Key Pattern file

1) Start
2) Authenticate user of public cloud pickup stego carrier from cloud storage
3) Authenticate user extract data from stego carrier
4) Authenticate user decompress extracted data with the help of compression pattern file.
5) Stop

*1) **Data extraction**:* Data extraction is exactly similar with data hiding process due to logical XOR operation. Only difference between algorithm of data hiding and extraction is extraction of amount of secret data bits. Extraction algorithm extracts only those number of bits which are hidden by sender. Time complexity of data extraction is directly proportional to amount of bits extracted. Algorithm 6 represents to data extraction steps performed by authenticated receiver. Efficiency of data extraction is measured by amount of data extracted and amount of data hidden. Author suggested an approach that is flexible with any data hiding mechanism. Author called data hiding approach as a virtual because of extremely minimum Q_e. Hidden data maybe an image, audio, video or text. There is no restriction on format of secret data and that's why it's possible to convert extracted data into the same format of hidden data. Comparative analysis which is mention in data extraction section of result analysis indicates that there is no single loss of character neither in it's format nor in placing.

V. RESULT ANALYSIS

A. Carrier space mapping

Consider an Image as Carrier with Size = $m * n$ Total_Pixels = $m * n$ = $m * n$ Total_Sample = $m * n * 3 = 3 * m * n$ Actual Space Required for Data Concealing = $\approx \frac{1}{3*m*n}$ Let m=100 and n=100 Actual_Space = $\frac{1}{100*100} = 0.0001 * 100 = 0.01$ % Unused Space = $0.9999 \approx 99.99$ % .It Means even though 99.99 % area of carrier on cloud not disturbed

B. Sample flipping

Any algorithm is incomplete without samples flipping except Least Significant Bit (LSB). Flipping is a process that minimizes Q_e and hence help in achieving audiovisual perceptual quality of carrier. Table IV shows Q_e with different bits flipping. Q_e goes on increasing from LSB to MSB position. More

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2021 doi:10.20944/preprints202111.0410.v1

TABLE V
DATA HIDING WITH SAMPLE FLIPPING

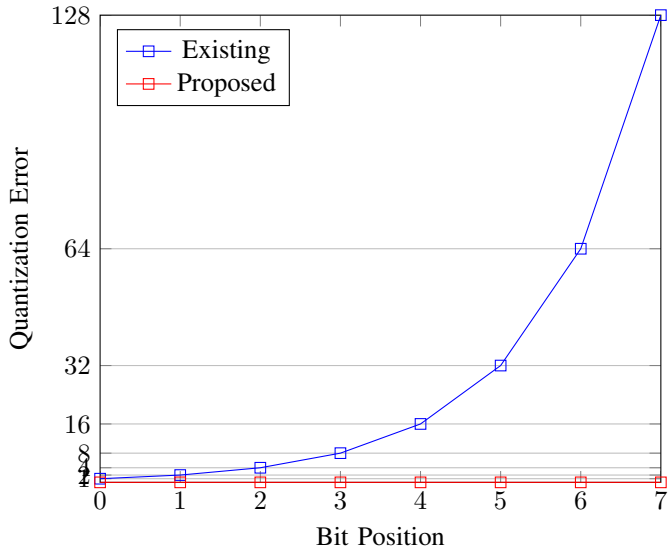
Original_Binary	Bit	Result_Binary	QE without Flipping	Flipping	QE with Flipping
11010100	1	11010101	1	11010101	1
11010100	1	11010110	2	11010011	1
11010100	1	11010100	0	No Need	0
11010100	1	11011100	8	11011111	1
11010100	0	11010100	0	No Need	0
11010100	0	11010100	0	No Need	0
11010100	0	11010000	4	11010011	-1
11010100	0	11010100	0	No Need	0

Algorithm 6 Data Extraction

Require: Stego Carrier, Compressed key Pattern file

1) Start
2) Input Carrier (Cr).
3) Extract Features (In Case of Image as a Carrier A, R, G, B).
4) Input Ch = No of Characters for Extractions
5) For I=0: Width (SC_r)
   For J=0: Height (SC_r)
   RS = RS + (Char)(A (I,J)XOR 255)
   If(Length (RS) >= Ch) ;
   Break;
   End
   End
   End
6) Save RS
7) Stop

flipping of sample bits required for MSB replacement. LSB to MSB movement for secret bit hiding generates an error that disturb originality of carrier and hence also violets properties of steganography. Q_e minimization through flipping of course a good task but complexity and flipping time are real challenge factors. Many of the authors/ researchers not suggested to use single steganography approach for a data hiding but they suggest to combine steganography with other approaches like cryptography and compression.



From the above graph it is observed that, the Q_e introduces due to existing flipping techniques, generates an error that exponentially rises with it's position. But with the proposed approach, Q_e is always zero because of virtual data hiding in carrier object. Some authors hides group of bits like 2,3,4 etc. Instead of hiding single bit. Hiding any number of bits will create an Q_e irrespective of it's position. From security point of view, Q_e is not affordable and not a good sign for sensitive data transmission.

C. *Data hiding*

These are the standard parameters considered as a difference measures between original and resultant carrier object. Many of these researchers did analysis on image processing using these parameters. average red $Avg_{red} = (\sum_{ri=1}^{n} P_{ri})/n$, average green $Avg_{green} = (\sum_{gi=1}^{n} P_{gi})/n$, and average blue $Avg_{blue} = (\sum_{bi=1}^{n} P_{bi})/n$ is the some all red, green &blue component divided by total number of red, green & blue component respectively in input or result object. Table VI indicate that, the difference in red, green and blue component of original carrier is minimum with stego carrier. Whereas MSE and PSNR is almost closer to 0 and 100 respectively. MSE and PSNR are inversely proportional to each other and hence indicate similarity between two comparative images.

Parameters	Difference/Value
Mean Square Error	0.01
Peak Signal to Noise	99.09
Min_Diff	1
Max_Diff	0
Average_Diff	0.11
Normalized Absolute Error (NAE)	0
Cross Co-Relation(CC)	1
Structural Content (SC)	1

TABLE VI
ORIGINAL VS STEGO IMAGE

Parameter	Original_Image	Result_Image
Red_Mean	103	103
Green_Mean	108	108
Blue_Mean	80	80
Mean	2440	2437
Pure_Height	5	5
Pure_Width	5	5
Entropy	443.75	443.75

TABLE VII
INDIVIDUAL IMAGE PARAMETERS

MSE $[= \frac{1}{n}\sum_{i=1}^{n}(c_i - c_i')^2]$ and PSNR $= 10 * \log_{10}\left(\frac{MAX^2}{MSE}\right)$ are used to compare images and their compression quality. MSE is the representation of cumulative square error between original

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 November 2021 doi:10.20944/preprints202111.0410.v1

and result image whereas PSNR represent to measure of peak difference between two images. Min difference $= Min\,|c_o - c_r|$ is the minimum value in array of difference between two images whereas Max difference $= Max\,|c_o - c_r|$ is the maximum value in array of difference between images. normalized absolute error (NAE) = $\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}|co_{ij}-cr_{ij}|}{\sum_{i=1}^{n}(co_{ij})}$ is used to compare images with different scale. It is the total absolute error normalized by error simply predicting the average of actual values. cross-correlation (CC) $\Phi_{xy} = \int_{-\infty}^{\infty} x(\tau - t)y(\tau)dt$ is the major of of two or more components of images with respect to each other. It is used to compare multiple data series pattern and find how they match up with each other at the best match point. structural content = $\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(co_{ij})^2}{\sum_{i=1}^{n}\sum_{j=1}^{m}(cr_{ij})^2}$ is also called as information about information and it provide information about the content type. Structural content helps to identify the distortion in result stego object after data hiding. $Entropy = \sum kP_k \log_2(P_k)$ is the measure of number of bits required to perform encryption on image data. Higher the value of Entropy indicates the more detail the image will be. It's statistical measure of randomness that can be used to characterize the texture of an image.

Technique	Embedding Capacity
Expansion	0.4258
Histogram Shifting	0.5687
Code Division Multiplexing	0.8654
Compression	0.7258
Contrast Enhancement	0.5578
Encrypted Domain	0.6952
Proposed	1

TABLE VIII
DATA HIDING CAPACITY COMPARISONS

Table VIII shows data hiding capacity of various algorithms compared with the proposed approach [8]. From the analysis, none of the techniques have 100% utilization of carrier for hiding data except propose technique. Existing compression techniques that have good data embedding capacity because of nature of compression of secret text but it not comparable with proposed approach. Proposed approach provides virtual data hiding capability that hides data virtually without any interfere of environmental noise.

D. *Data extraction*

Performance of the entire proposed system is based on how accurately data extracted from stego object. This stego object is accessible to the public when it is kept on public cloud storage. Extracted data is always in compressed format which is decompress with the help of key replacement file and the result obtained is compared with original hidden text. Author have developed a model that compare hidden text and extracted text and found that there is no mismatch or misspelled of character.

VI. Conclusion

Proposed approach good fits for an applications where there is need to transfer huge amount of data (sensitive) in a single stroke. Intruder will never get a clue about existence of data because of nature of virtual hiding of secret bits. Compression algorithms does well in both the level. Finally, only 2 to 3 characters are the output of compression level 2 which is the best result forever. Additional products (compressed pattern key file) generated which needs to be taken care while sharing with authenticated receiver. Any leakage in sharing compress pattern key file will put an intruder into an observation of public cloud storage objects. Sender or original author almost relax about presence of sensitivity data behind carrier. Even though carrier stego object corrupt, lost or modified but it's not possible to touch data behind it

VII. Future Scope

Future scope are vision improvement for any propose concept. Every project have it's positive as well as negative aspects which varies from user to user. Proposed approach have some limitations which should be considered as future scope. The additional compress key pattern file generated needs careful attention while transfer from sender to authenticate sources. Compression of proposed approach is depend on patterns which are used during the process. If these key patterns leaks somewhere, it will be easy for an intruder to identify data. Storage and transmission of compress key pattern file is really a challenging tasks over unsecured wireless network. Author always focus on problem statement like server crash, integrity, backup crash and others but still 100% resolution of these statements not possible impractical era.

References

[1] M. Nosrati, R. Karimi, H. Nosrati, and A. Nosrati, Embedding stegotext in cover images using linked list concepts and LSB technique, Journal of American Science, Vol. 7, No. 6, 2011, pp. 97-100.

[2] Wen-Chung Kuo, Dong-Jin Jiang, Yu-Chih Huang, A Reversible Data Hiding Scheme Based on Block Division, Congress on Image and Signal Processing, Vol. 1, 27-30 May 2008, pp. 365-369

[3] Soumyendu Das, Subhendu Das, Bijoy Bandyopadhyay, Sugata Sanyal,Steganography and Steganalysis: Different Approaches, 35 International Journal of Computers, Information Technology and Engineering (IJCITAE), Vol. 2, No 1, June, 2008, Serial Publications, pp. 1-11

[4] Ming Sun Fu and O.C. Au, Data hiding watermarking for halftone images", IEEE Transactions on Image Processing, Vol.11, No. 4, Apr. 2002, pp.477-484.

[5] H. B. Kekre, Archana Athawale, Archana Athawale, Uttara Athawale, Information Hiding in Audio Signals, International Journal of Computer Applications IJCA, Vol. 7, No. 9, Foundation of Computer Science, New York, USA, pp. 14-19 2010.

[6] Xiaoyin Qi, Xiaoni Li, Mianshu Chen, Hexin Chen, Research on CAVLC audiovideo synchronization coding approach based on H.264, IEEE International Conference on Uncertainty Reasoning and Knowledge Engineering (URKE), Vol. 2, 4- 7 Aug. 2011, pp.123-126

[7] T. Hong, W. Chen and H. Wu, An improved reversible data hiding in encrypted images using side match, IEEE Signal Processing Lett., vol.19, no. 4, pp. 199-202, 2012.

[8] Asha Jose,Kamalraj Subramaniam, Comparative analysis of reversible data hiding schemes, IET Journal,March 2020 https://doi.org/10.1049/iet-ipr.2019.1066

[9] X. Zhang, N. Wuwong, H. Li, and X. Zhang, "Information Security Risk Management Framework for the Cloud Computing Environments," 2019, pp. 1328–1334

[10] F. Xie, Y. Peng, W. Zhao, D. Chen, X. Wang, and X. Huo, "A risk management framework for cloud computing," in Cloud Computing and Intelligent Systems (CCIS), 2012 IEEE 2nd International Conference on, 2018, vol. 1, pp. 476–480

[11] S. Tanimoto, M. Hiramoto, M. Iwashita, H. Sato, and A. Kanai, "Risk Management on the Security Problem in Cloud Computing," in Computers, Networks, Systems and Industrial Engineering (CNSI), 2011 First ACIS/JNU International Conference, pp.147,152.

[12] M. A. Alhomidi and M. J. Reed, "Security risk analysis as a service," in Internet Technology and Secured Transactions (ICITST), 2013 8th International Conference for, 2017, pp. 156–161.

[13] H. Takabi, J. B. D. Joshi, and G.-J. Ahn, "Secure Cloud: Towards a Comprehensive Security Framework for Cloud Computing Environments," Computer Software and Applications Conference Workshops (COMPSACW), 2019 IEEE 34th Annual, pp. 393–398.

[14] F. Xie, Y. Peng, W. Zhao, D. Chen, X. Wang, and X. Huo, “A risk management framework for cloud computing,” in Cloud Computing and Intelligent Systems (CCIS), 2012 IEEE 2nd International Conference on, 2018, vol. 1, pp. 476–480.

[15] Mehdi Hussain, Ainuddin Wahid Abdul Wahab, Yamani Idna Bin Idris, Image Steganography in spatial domain: A survey, Signal Processing:Image Communication, vol. 65, pp. 46-66, 2018

[16] Khan Muhammad, Jamil Ahmad, Seungmin Rho, Sung Wook Baik,Image Steganography for authenticity of visual contents in social networks, Multimed Tools Appl, vol. 76, pp. 18985-19004, 2017.

[17] G. Smitha, E. Baburaj. A survey on image steganography based on block-based edge adaptive based on Least Significant Bit Matched Revisited (LSBMR) algorithm. Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 132-139, 2016

[18] Jiantao Zhou, Weiwei Sun, Li Dong, Xianming Liu, Oscar C. Au Yuan Yan Tang,Secure Reversible Image Data Hiding over Encrypted Domain via Key Modulation, IEEE Trans. On Circuits and Systems for Video Technology, vol. 26, issue 3, pp.441-452,2015.

[19] X. Zhang, Z. Qian, G. Feng, and Y. Ren,Efficient reversible data hiding in encrypted images, J. Vis. Commun. Image R., vol. 25, no. 2, pp. 322- 328, 2014.

[20] C. Qin, C. -C. Chang, Y.-H. Huang, and L.-T. Liao, An inpaintingAssisted reversible steganographic scheme using a histogram shifting mechanism, IEEE Trans. Circuits Syst . Video Technol., vol. 23, no. 7, pp. 1109-1118, 2013