International Journal on Emerging Technologies (Special Issue on ICRIET-2016) **7**(2): 338-340(2016)

ISSN No. (Print) : 0975-8364
ISSN No. (Online) : 2249-3255

A General Study on Secure Data Compression

Nagesh Salimath [*] ***and Dr.Jitendra Sheetlani*** [**]
[*]*Research Scholar, Computer Science*
Sri Satya Sai University of Technology & Medical Sciences, Sehore, (Madhya Pradesh), India
[**]*Dean & Associate Professor, School of Computer Application*
Sri Satya Sai University of Technology & Medical Sciences, Sehore, (Madhya Pradesh), India

(Corresponding author: Nagesh Salimath)
(Received 28 September, 2016 Accepted 29 October, 2016)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: **This paper describes the importance of data compression techniques for networks and cloud storage. The data is compressed in several ways one has to choose which method is suitable for him and the compression depends upon the type of file also. In current trend data plays very significant role in rapid growth of the technology era, the users are in threat for their valuable data and so on. Regarding compression in data, the signal processing, compression of data, source coding or reduction in bit-rate convolves the encoding of information utilizing smaller number of bits than the real presentation. In the compression there has been two kinds those are lossy or lossless. The lossless compression is eliminating the statistical redundancy. The loss of information is not present in this technique, and the lossy compression reducing the bits with deleting the unwanted or non-significant data. The mechanism in which the reduction of data size takes place is called the data compression. In the compression algorithms are measured in two aspects which are one complexity and amount of compression. The main advantage of data compression is to make speed transmission of data from once place to another in short time with security also. For perform the transmission smooth in network one has to ensure the data compression ration so, that the data has to move fast in networks and to utilize the minimum storage in cloud also it is helpful to the user to give less amount for less storage**.

Keywords: Data Compression, Networks, Cloud Storage.

I. INTRODUCTION

The data compression, is more important in today's rapidly growing technology, because users of technology are putting so many things on internet. To facilitate the user to exchange the information, the internet technology should equip with more facilities, like giving speed internet, space, applications and so on. In that another main thing plays a vital role that is data compression, because users are sending and receiving heavy information in the form of video, audio and images by the social networks like facebook, twitter and etc. Where large volume of data is moving from one location to another location to make this possible the developers should develop a compression technology to the users. Now a days so many users are share their personnel information on internet and they expect their information should be confidential and secure, to provide this feature the developers of the technology look for secured data compression where data is compressed in secured manner. The data should compress to reduce the memory and speed in transmitting that information. To make this the data should be compressed in such a manner that Can remove the so much of redundancies of data stored and it cuts the cost of storage. Many data stored in the disk now a days should few statistical redundancy and is easy compressed. Shrinking the information removes the size of storage in physical required to as less as in 30% the space. The customers are considering a ways to leverage the lossless and transparent compression in storage medium to controlling the degree of storage.

In this paper some data compression technologies are discussed that will give a glimpse to a reader to understand and usage of good compression technology. Compression may be lossless or lossy. The lossless compression decrease the bits by identifying and removing the quantitative redundancy. There is zero percent of loss in lossless compression. The lossy compression decreases the bits by eliminating the unwanted or least important information. The steps of minimizing the size of a data file is concerned as compression of data.

In this regard the transmission of data, is called source coding (the data stored or else transmitted it has to encode first) in bilateral channel coding. The reduction of file size is important in compressing the resources need to transmit and store the data. Calculation of resources are absorbed in the process of compression and commonly, in the bidirectional of the process (decompression). The compression of data is referenced to a space and time complexity trades. For example, a scheme of compression for video may need the costly hardware parts for the video has to be decompressed fast enough to viewed as it is become decompressed, and the option to decompress the video before it watching it should be not clear or it may need the extra storage. The structure of compression in data it include trade-offs in various facts, including the size of compression, the size of distortion is introduced and calculation of resources need to compress and decompress the data

II. DATA COMPRESSION LOSSY

The compression of data in lossy works in very dynamically. The programs are easily deletes the " Unwanted" bits of information. Joining the files so the it become lesser. This kind of compression is utilized a lot of reduction of the size of files of bitmap format images, which are heavier to store. To observe the works, now consider the computer is going to compress a scanned picture. Lossy compression is not reversible compression is the case of encoding the data procedure that employs the inaccurate approximations and partial data eliminated to present the content.

The lossy compression of file results in the loosing the original data and quality it oftenly took the more less space in disk compare to original one. For instance, a JPEG but also be utilized in audio file formats, for example MP3s or AAC files. The "lossyness" in image file should show the random edges or pixilated areas. The audio, the compression in lossyness should produce a watery sound or decrease the different range of the sound.

Due to compression in lossy eliminates the data from the main file, it is resulting the file often took the less space in disk compare to main file. For instance, a JPEG picture may decrease the file size of imag's with 80% more, with less observable effect. As same as, a MP3 compressed file should be ten percent size of the original audio file it may sound much same as original.

In this lossy compression JPEG and MP3 compression are both eliminates the data from the main file,

which should be considerable on the close examination. In the both of these size reduction algorithms may also the different kinds of " setting in quality of file", which gives the method of compressed the file. The quality includes the trade-off between file size quality. File that used the larger compression will take the less space, but may not see the sound as good as a less file compress. The images are in audio allow the compression in lossless, which not reduce the files quality.

Here are the some lossy compression methods, wavelet compression, the lossy compression of an image by converting it into a set of mathematical expressions, JPEG (Joint Photographic Experts Group), this compression method uses the discrete cosine transform (DCT). DCT is one of the signal processing technique, thus is uses the mathematical operation which the frame/field converts the source of video from the spatial domain converts to the frequency domain. The domain transform, the steps of reduction of information is called the quantization. It is the method of optimal reduction in a large size of scale into the smaller the transform-domain, is a easier presentation of the image due to high-frequency co-efficients, which gives decrease the overall image than other coefficients.

III. LOSSLESS DATA COMPRESSION

Lossless compression is a class of data compression algorithms that allows the original data to be perfectly reconstructed from the compressed data. By contrast, lossy compression permits reconstruction only of an approximation of the original data, though this usually improves compression rates (and therefore reduces file sizes). Lossless compression is used in cases where it is important that the original and the decompressed data be identical, or where deviations from the original data could be deleterious. Typical examples are executable programs, text documents, and source code. Some image file formats, like PNG or GIF, use only lossless compression.

If, we see the history of lossless compression technologies we got, Data compression has only played a significant role in computing since the 1970s, when the Internet was becoming more popular and the Lempel-Ziv algorithms were invented, but it has a much longer history outside of computing. Morse code, invented in 1838, is the earliest instance of data compression in that the most common letters in the English language such as "e" and "t" are given shorter Morse codes. Later, as mainframe computers were starting to take hold in 1949, Claude Shannon and Robert Fano invented Shannon-Fano coding. Their algorithm assigns codes to symbols in a given block of data based on the probability of the symbol occurring. The probability is of a symbol occurring is inversely proportional to the length of the code, resulting in a shorter way to represent the data. In previous implementations of Shannon-Fano and Huffman coding are employed the hardware and more codes.

It was not there till 1970s and the development of the Internet and Online storage the software compression is developed that uses the Huffman codes for in dynamically generated the input data. Next, in 1977, Abraham Lempel and Jacob Ziv presented the paper that given the ground braking LZ77 algorithm, the very first algorithm to utilize a dictionary for compress the data. Most specifically, LZ77 utilized a dynamic dictionary oftenly called a sliding window. In the year 1978, the both presented their paper on LZ78 algorithm which used a dictionary like LZ77 uses, this method converts the input data and generates a static dictionary rather than generating it dynamically. There were some examples in lossless compression technology, the Run-Length encoding is a very simple compression where it inter-changing the runs of two or more same character with the number of which represents the length of the run, allowed the original character; the characters are coded as runs of 1. RLE is used for highly-redundant data, images in indexed with many pixels of color in a row, or in join with other compression methods like the Burrows-Wheeler Transform. In the Burrows-Wheeler Transform it is a compression scheme developed in 1994 that has aimed to reversible transform a block of input data in the amount of runes of identical characters are maximized. In the self BWT not going to perform any compression techniques, it is simple changes the input and it can easily coded by Run-Length Encoder or other secondary compression technique.

The Encoding in the Entropy, the data compression means the smallest number of bits needed, to present a literal. A fundamental entropy coder joins a statistical model and a coder. The file which is fed as input parsed to used to create a statistical prototype that consists of the probabilities of a given literal appearing. The coder will utilize the model of statistical to determine the bit codes to assign a each literal in most symbols have the shortest codes and the less common symbols got the shortest codes and the less common literals have the longest codes. Shannon-Fano coding, one the earliest compression methods, developed in the year of 1949 by Claude Shannon and Robert Fano. This kind of technique includes the generating a binary tree to represent the probabilities of each symbol coming. Huffman coding, is the another flavor of top down to create the nearest result. Arithmetic coding : this method has implemented in the year 1979 at IBM, which was investigating the data compression methods for the use in the mainframes. Arithmetic coding is the best optimal entropy coding method if the objective is the good reduction ratio since it usually achieves good results as compare to Huffman Coding. However the less complications compared with other coding techniques, like numerous lossless compression methods we have so many.

CONCLUSION

This paper thrown a light on data compression, it has got some informative This paper discusses the different type of lossy and lossless data compression. And different type of research paper on the based on data compression. This paper only discusses the general idea of data compression. Today, many compression techniques are developed and some techniques are in process .But this paper only discusses the general idea about the Shannon-Fano, Huffman coding, arithmetic coding. Better manner and relate it to the future work.

REFERENCES

[1] Amandeep Singh Sidhu, "Research Paper on Text Data Compression Algorithm using Hybrid Approach " , *IJCSMC*, Vol. **3**, Issue. 12, pg.01 – 10, 2014.

[2] Neha Sharma, Jasmeet Kaur, Navmeet Kaur, " A Review on various Lossless Text Data Compression Techniques" , ijoes, 2014.

[3] Brar et al., "A Survey on Different Compression Techniques and Bit Reduction Algorithm for Compression of Text/Lossless Data" ,*ijarcsse,* vol **3**, issue 3, pp 579-582, 2013.

[4] Amarjit Kaur et al., "A Review on Data Compression Techniques ", vol **5**, issue 1, pp 769-773, 2015.

[5] I Made Agus Dwi Suarjaya, "A New Algorithm for Data Compression Optimization", Vol. **3**, No.8, 2012.

[6] Saumya mishra, Shraddha sing, "A Survey Paper on Different Data Compression Techniques" , ijoar, vol 6, issue 5, pp 738-740, 2016.

[7] http://ethw.org / History_of_Lossless_Data_Compression_Algorithms

[8] https: //en.wikipedia.org /wiki/Lossless_compression#Text_and_image

[9] https://en.wikipedia.org/wiki/Lossy_compression

[10] Jacob Ziv et al, " A Universal Algorithm for Sequential Data Compression ", IEEE Transaction on Information Theory, Vol **23**, pp 337- 343, 1977.